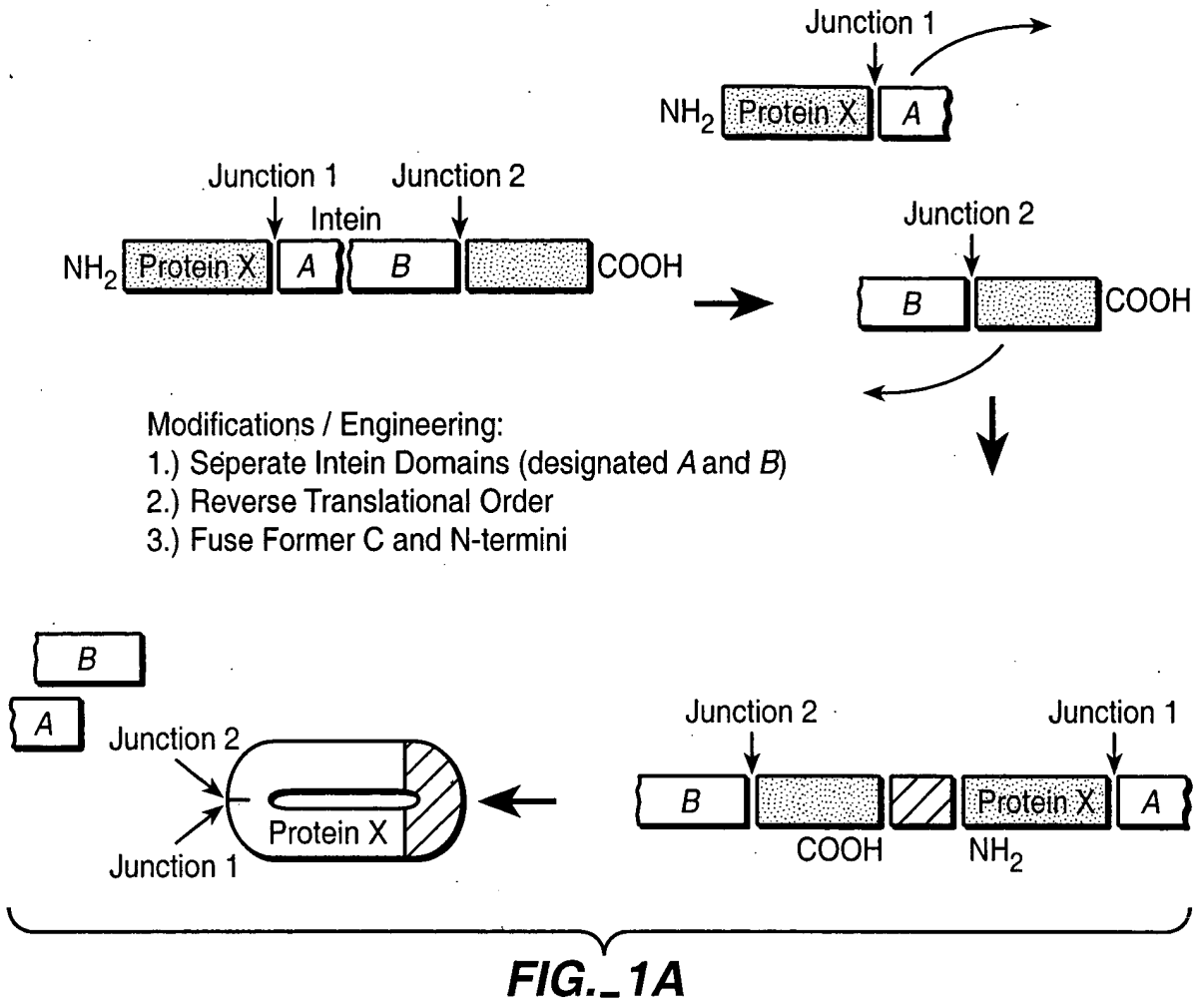

Junction 1
NH2 Protein X A
Junction 1 Junction 2
Intein
NH2 Protein X A B COOH
Junction 2
B COOH
Modifications / Engineering:
1.) Seperate Intein Domains (designated A and B)
2.) Reverse Translational Order
3.) Fuse Former C and N-termini
B
A
Junction 2
Protein X
Junction 1
Junction 2
Junction 1
B
Protein X
A
COOH
NH2

FIG._1A

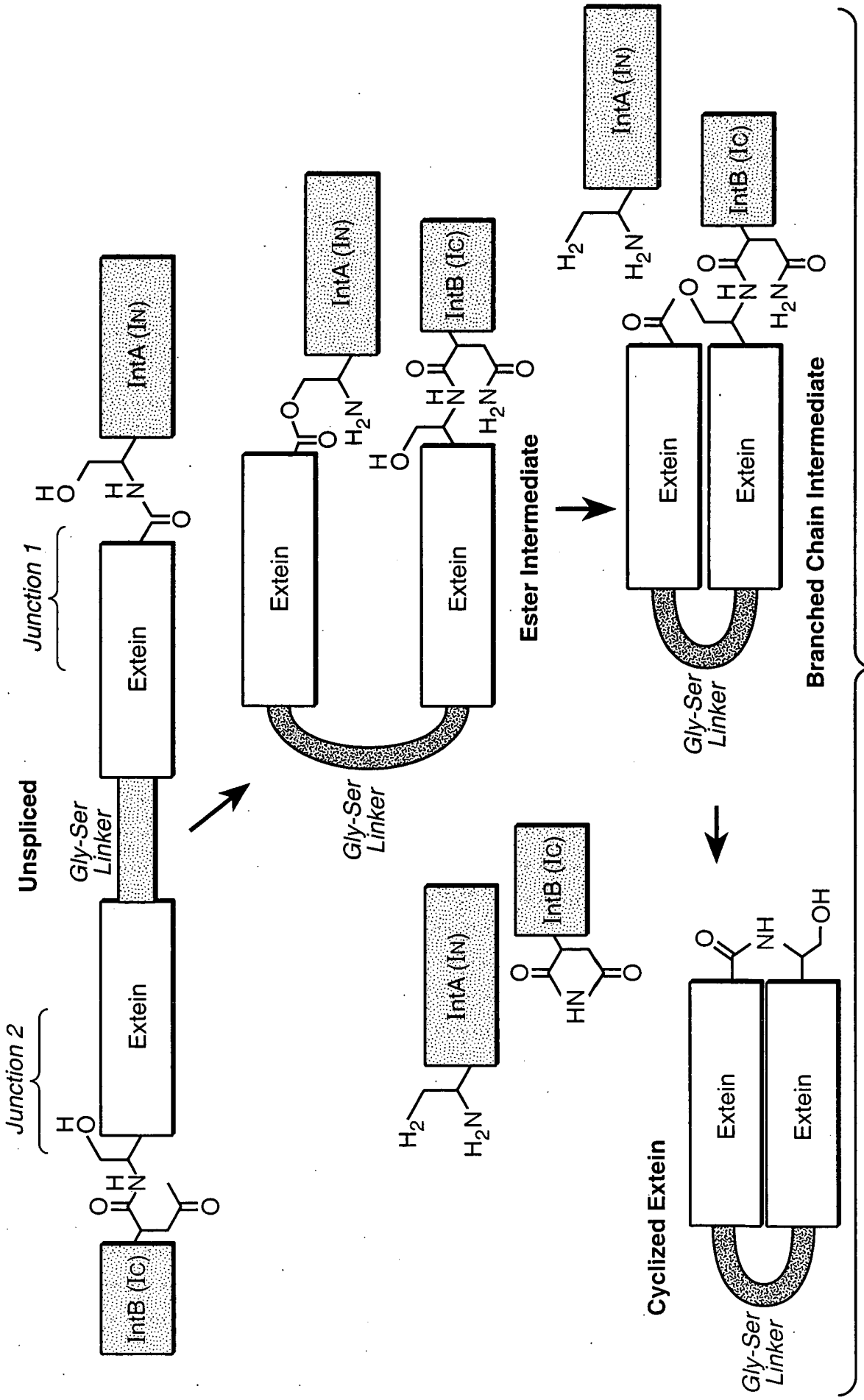

IntB (Ic)
IntA (In)
Extein
Extein
Junction 2
Junction 1
Gly-Ser Linker
Unspliced
Extein
Extein
IntA (In)
IntB (Ic)
Gly-Ser Linker
Ester Intermediate
IntA (In)
IntB (Ic)
Extein
Extein
Gly-Ser Linker
Branched Chain Intermediate
IntB (Ic)
IntA (In)
Extein
Extein
Gly-Ser Linker
Cyclized Extein

FIG._1B

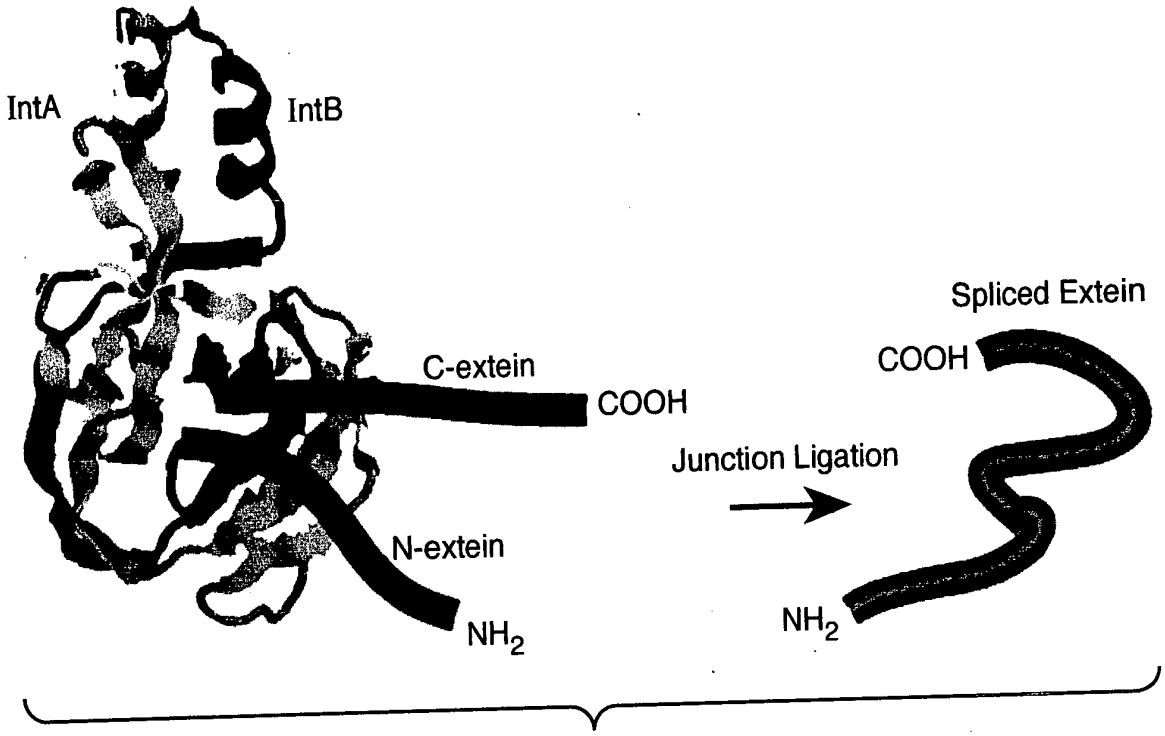

IntA
IntB
C-extein
COOH
N-extein
NH_2
Junction Ligation
Spliced Extein
COOH
NH_2

FIG._2A

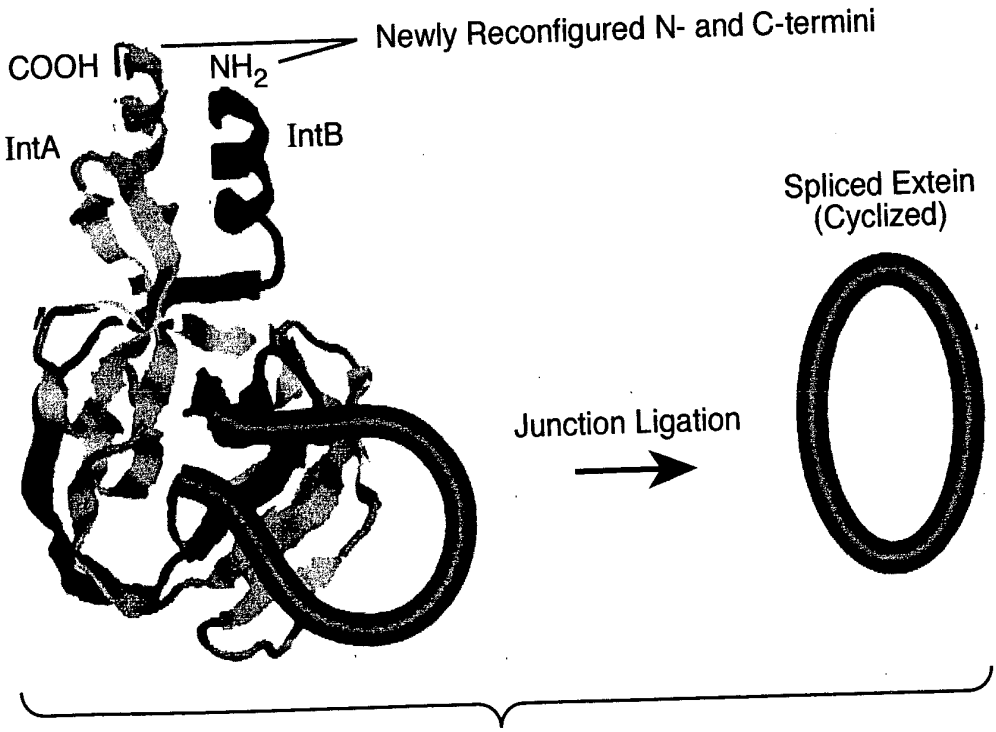

Newly Reconfigured N- and C-termini
COOH
NH_2
IntA
IntB
Junction Ligation
Spliced Extein
(Cyclized)

FIG._2B

GCISGDSLISLASTGKRVSIKDLLDEKDFEIWAINEQTMKLESAKVSRVFCTGKKLVYILKT
RLGRTIKATANHRFLTIDGWKRLDELSLKEHIALPRKLESSSLQLMSDEELGLLGHLIGDGC
TLPRHAIQYTSNKIELAEKVVELAKAVFGDQINPRISQERQWYQVYIPASYRLTHNKKNPIT
KWLENLDVFGLRSYEKFVPNQVFEQPQRAIAIFLRHLWSTDGCVKLIVEKSSRPVAYYATSS
EKLAKDVQSLLLKLGINARLSKISQNGKGRDNYHVTITGQADLQIFVDQIGAVDKDKQASVE
EIKTHIAQHQANTNRDVIPKQIWKTYVLPQIQIKGITTRDLQMRLGNAYCGTALYKHNLSRE
RAAKIATITQSPEIEKLSQSDIYWDSIVSITETGVEEVFDLTVPGPHNFVANDIIVHNS

FIG._3A

YCITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRHGNPVLADRLFHSGEHPVYTV
RTVEGLRVTGTANHPLLCLVDVAGVPTLLWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPE
FAPTTYTVGVPGLVRFLEAHHRDPDAQAIADELTDGRFYYAKVASVTDAGVQPVYSLRVDTA
DHAFITNGFVSHNT

FIG._3B

ECLTSDHTVLTTRGWIPIADVTLDDKVAVLDNNTGEMSYQNPQKVHKYDYEGPMYEVKTAGV
DLFVTPNHRMYVNTTNNTTNQNYNLVEASSIFGKKVRYKNDAIWNKTDYQFILPETATLTGH
TNKISSTPAIQPEMNAWLTFFGLWIANGHTTKIAEKTAENNQQKQRYKVILTQVKEDVCDII
EQTLNKLGFNFIRSGKDYTIENKQLWSYLNPFDNGALNKYLPDWVWELSSQQCKILLNSLCL
GNCLFTKNDDTLHYFSTSERFANDVSRLALHAGTTSTIQLEAAPSNLYDTIIGLPVEVNTTL
WRVIINQSSFYSYSTDKSSALNLSNNVACYVNAQSALTLEQNSQKINKNTLVLTKNNVKSQT
MHSQRAERVDTALLTQKELDNSLNHEILINKNPGTSQLECVVNPEVNNTSTNDRFVYYKGPV
YCLTGPNNVFYVQRNGKAVWTGNS

FIG._3C

LCVAPETMILTEDGQFPIKDLEGKIIKVWNGNEFSSVTVVKTGTEKELLEVELSNGCTLSCT
PEHKFIIVKSYTEAKKQKTDDNAIANAERVDAQDLKPRMKLIKFDLPTLFGNSEHDIKYPYT
HGFFCGDGTYTKYGKPQLSLYGDKKELLTYLDVRTMTGLEDASGRLNTWLPLDLAPKFDVPI
NSSLECRMEWLAGYLDADGCVFRNGTNESIQVSCIHLDFLKRIQLLLIGMGVTSKITKLHDE
KITTMPDGKGGQKPYSCKPIWRLFISSSGLYHLSEQGFETRRLKWEPRQPQRNAERFVEVLK
VNKTGRVDDTYCFTEPINHAGVFNGILTGQC

FIG._3D

GCFTKGTQVMMADGADKSIESIEVGDKVMGKDGMPREVVGLPRGYDDMYKVRQLSSTRRNAK
SEGLMDFTVSADHKLILKTKQDVKIATRKIGGNTYTGVTFYVLEKTKTGIELVKAKTKVFGH
HIHGQNGAEEKAATFAAGIDSKEYIDWIIEARDYVQVDEIVKTSTTQMINPVHFESGKLGNW
LHEHKQNKSLAPQLGYLLGTWAGIGNVKSSAFTMNSKDDVKLATRIMNYSSKLGMTCSSTES
GELNVAENEEEFFNNLGAEKDEAGDFTFDEFTDAMDELTINVHGAAASKKNNLLWNALKSLG
FRAKSTDIVKSIPQHIAVDDIVVRESLIAGLVDAAGNVETKSNGSIEAVVRTSFRHVARGLV
KIAHSLGIESSINIKDTHIDAAGVRQEFACIVNLTGAPLAGVLSKCALARNQTPVVKFTRDP
VLFNFDLIKSAKENYYGITLAEETDHQFLLSNMALVHNC

FIG._3E

GCLSYATNQPYFLKSDNVNFSKLTSLKVSNHYILSATLELLIPFQYNRIYPIVSLIKRELQT
GYKVVYELDFYISVIVSTVEHYVLTLNGWKRILELTVDDLVATLDIQYLIYNNTEVDLFSSN
VIFSSVINLICMNRINVYDFWIPKTNNFFVNALLVHNS

FIG._3F

GCISKFSHIMWSHVSKPLFNFSIKKSHMHNFNKNIYQLLDQGEAFISRQDKKTTYKIRTNSE
KYLELTSNHKILTLRGWQRCDQLLCNDMITTQIGFELSRKKKYLLNCIPFSLCNFETLANIN
ISNFQNVFDFAANPIPNFIANNIIVHNS

FIG._3G

GCFAKGTNVLMADGSIECIENIEVGNKVMGKDGRPREVIKLPRGRETMYSVVQKSQHRAHKS
DSSREVPELLKFTCNATHELVVRTPRSVRRLSRTIKGVEYFEVITFEMGQKKAPDGRIVELV
KEVSKSYPISEGPERANELVESYRKASNKAYFEWTIEARDLSLLGSHVRKATYQTYAPILYE
NDHFFDYMQKSKFHLTIEGPKVLAYLLGLWIGDGLSDRATFSVDSRDTSLMERVTEYAEKLN
LCAEYKDRKEPQVAKTVNLYSKVVRGNGIRNNLNTENPLWDAIVGLGFLKDGVKNIPSFLST
DNIGTRETFLAGLIDSDGYVTDEHGIKATIKTIHTSVRDGLVSLARSLGLVVSVNAEPAKVD
MNGTKHKISYAIYMSGGDVLLNVLSKCAGSKKFRPAPAAAFARECRGFYFELQELKEDDYYG
ITLSDDSDHQFLLANQVVVHNC

FIG._3H

GCFAYGTRGALADGTTEKIGKIVNQKMDVEVMSYDPDTDQVVPRKVVNWFNNGPAEQFLQFT
VEKSGGNGKSQFAATPNHLIRTPAGWTEAGDLVAGDRVMAAEPHRLSDQQFQVVLGSLMGDG
NLSPNRRDRNGVRFRMGHGAKQVDYLQWKTALLGNIKHSTHVNDKGATFVDFTPLPELAELQ
RAVYLGDGKKFLSEENFKALTPLALVFWYMDDGPFTVRSKGLQERTAGGSGRIEICVEAMSE
GNRIRLRDYLRDTHGLDVRLRLSGAAGKSVLVFSTASSAKFQELVAPYITPSMEYKLLPRFR
GQGAVTPQFVEPTQRLVPARVLDVHVKPHTRSMNRFDIEVEGNHNYFVDGVMVHNS

FIG._3I

YCLSFGTEILTVEYGPLPIGKIVSEEINCSVYSVDPEGRVYTQAIAQWHDRGEQEVLEYELE
DGSVIRATSDHRFLTTDYQLLAIEEIFARQLDLLTLENIKQTEEALDNHRLPFPLLDAGTIK

FIG._3J

KALALDTPLPTPTGWTAMGDVAVGDELLAVDEAPTRVVAATEVMLGRPCYEIEFSDGTVIVA
DAQHQWPTSYGIRTSAQLRCGLDIIAAAGSTPRHAGRLTTAAFMAPVLCIDSVRRVRSVPVR
CVEVDNAAHLYLAGRGMVPTHNS

FIG._3K

GALAYDEPIYLSDGNIINIGEFVDKFFKKYKNSIKKEDNGFGWIDIGNENIYIKSFNKLSLI
IEDKRILRVWRKKYSGKLIKITTKNRREITLTHDHPVYISKTGEVLEINAEMVKVGDYIYIP
KNNTINLDEVIKVETVDYNGHIYDLTVEDNHTYIAGKNEGFAVSNC

FIG._3L

GALYDFSVIQLSNGRFVLIGDLVEELFKKYAEKIKTYKDLEYIELNEEDRFEVVSVSPDlKA
NKHVVSRVWRRKVREGEKLIRIKTRTGNEIILTRNHPLFAFSNGDVVRKEAEKLKVGDRVAV
MMRPPSPPQTKAVVDPAIYVKISDYYLVPNGKGMIKVPNDGIPPEKAQYLLSVNSYPVKLVR
EVDEKLSYLAGVILGDGYISSNGYYISATFDDEAYMDAFVSVVSDFIPNYVPSIRKNGDYTI
VTVGSKIFAEMLSRIFGIPRGRKSMWDIPDVVLSNDDLMRYFIAGLFDADGYVDENGPSIVL
VTKSETVARKIWYVLQRlGIISTVSRVKSRGFKEGELFRVIISGVEDLAKFAKFIPLRHSRK
RAKLMEILRTKKPYRGRRTYRVPISSDMIAPLRQMLGLTVAELSKLASYYAGEKVSESLIRH
IEKGRVKEIRRSTLKGIALALQQIAKDVGNEEAWVRAKRLQlIAEGDVYWDEVVSVEEVDPK
ELGIEYVYDLTVEDDHNYVANGILVSNC

FIG._3M

PCVSGDTIVMTSGGPRTVAELEGKPFTALIRGSGYPCPSGFFRTCERDVYDLRTREGHCLRL
THDHRVLVMDGGLEWRAAGELERGDRLVMDDAAGEFPALATFRGLRGAGRQDVYDATVYGAS
AFTANGFIVHNC

FIG._3N

GCIDGKAKIIFENEGEEHLTTMEEMYERYKHLGEFYDEEYNRWGIDVSNVPIYVKSFDPESK
RVVKGKVNVIWKYELGKDVTKYEIITNKGTKILTSPWHPFFVLTPDFKIVEKRADELKEGDI
LIGGMPDGEDYKFIFDYWLAGFIAGDGCFDKYHSHVKGHEYIYDRLRIYDYRIETFEIINDY
LEKTFGRKYSIQKDRNIYYIDIKARNITSHYLKLLEGIDNGIPPQILKEGKNAVLSFIAGLF
DAEGHVSNKPGIELGMVNKRLIEDVTHYLNALGIKARIREKLRKDGIDYVLHVEEYSSLLRF
YELIGKNLQNEEKREKLEKVLSNHKGGNFGLPLNFNAFKEWASEYGVEFKTNGSQTIAIIND
ERISLGQWHTRNRVSKAVLVKMLRKLYEATKDEEVKRMLHLIEGLEVVRHITTTNEPRTFYD
LTVENYQNYLAGENGMIFVHNT

FIG._3O

NSILPEEWVPLIKNGKVKIFRIGDFVDGLMKANQGKVKKTGDTEVLEVAGIHAFSFDRKSKK
ARVMAVKAVIRHRYSGNVYRIVLNSGRKITITEGHSLFVYRNGDLVEATGEDVKIGDLLAVP
RSVNLPEKRERLNIVELLLNLSPEETEDIILTIPVKGRKNFFKGMLRTLRWIFGEEKRVRTA
SRYLRHLENLGYIRLRKIGYDIIDKEGLEKYRTLYEKLVDVVRYNGNKREYLVEFNAVRDVI
SLMPEEELKEWRIGTRNGFRMGTFVDIDEDFAKLLGYYVSEGSARKWKNQTGGWSYTVRLYN
ENDEVLDDMEHLAKKFFGKVKRGKNYVEIPKKMAYIIFESLCGTLAENKRVPEVIFTSSKGV
RWAFLEGYFIGDGDVHPSKRVRLSTKSELLVNGLVLLLNSLGVSAIKLGYDSGVYRVYVNEE
LKFTEYRKKKNVYHSHIVPKDILKETFGKVFQKNISYKKFRELVENGKLDREKAKRIEWLLN
GDIVLDRVVEIKREYYDGYVYDLSVDEDENFLAGFGFLYAHNS

FIG._3P

DSVTGETEIIIKRNGKVEFVAIEELFQRVDYRIGEKEYCVLEGVEALTLDNRGRLVWKSVPY
VMRHRTNKRIYRVWFTNSWYLDVTEDHSLIGYMNTSKVKPGKPLKERLVEVKPGELGESVKS
LITPNRAIAHGIRVNPIAVKLWELIGLLVGDGNWGGQSNWAKYNVGLSLGLDKEEIEEKILK
PLKNTGIISNYYDKSKKGDVSILSKWLARFMVRYFKDESGSKRIPEFMFNLPREYIEAFLRG
LFSADGTVSLRKGVPEVRLTSVNPELSSSVRKLLWLVGVSNSMFVETNPNRYLGKESGTHSV
HVRIKDKHRFAERIGFLLDRKATKLSENLGGHTSKKRAYKYDFDLVYPKKVEEIAYDGYVYD
IEVEGTHRFFANGILVHNT

FIG._3Q

KCLLPEEKVVLPEIGLVTLRELFELANEVVVKDEEKEVRKLGKMLTGVDERGNVKLLNALYV
WRVAHKGEMIRVKVNGWYSVTVTPEHPFLTNRGWVKAGELKEGDYIAIPRRVYGNEDlMKFS
KIAKELGIKGDEKEFYLAGASlDIPIKVLFLAPSKLVSAFLRGYFDAKGVVRENYIEVPLFE
DLPLLlLRFGIVSRIEKSTLKISGKRNLELFRKHVGFTDSEKAKALDELISKAKESERYPIl
EELRRLGLLFGFTRNELRIEENPTYEVlMEILERIERGSPNLAEKIAVLEGRIKEENYLRIL
EEEGLIENGKLTELGKELLEVWRNREFDSKDVDYVRNIVENLVFLPVEKVERIEYEGYVYDV
TTETHNFVANGILVHNT

FIG._3R

QCFSGEEVIIVEKGKDRKVVKLREFVEDALKEPSGEGMDGDIKVTYKDLRGEDVRILTKDGF
VKLLYVNKREGKQKLRKIVNLDKDYWLAVTPDHKVFTSEGLKEAGEITEKDEIIRVPLVILD
GPKIASTYGEDGKFDDYIRWKKYYEKTGNGYKRAAKELNIKESTLRWWTQGAKPNSLKMIEE
LEKLNLLPLTSEDSRLEKVAIILGALFSDGNIDRNFNTLSFISSERKAIERFVETLKELFGE
FNYEIRDNHESLGKSILFRTWDRRIIRFFVALGAPVGNKTKVKLELPWWIKLKPSLFLAFMD
GLYSGDGSVPRFARYEEGIKFNGTFEIAQLTDDVEKKLPFFEEIAWYLSFFGIKAKVRVDKT
GDKYKVRLIFSQSIDNVLNFLEFIPISLSPAKREKFLREVESYLAAVPESSLAGRIEELREH
FNRIKKGERRSFIETWEVVNVTYNVTTETGNLLANGLFVKNS

FIG._3S

LCLTPDTYVVLGDGRIETIEDIVNAKERNVLSLDLDNLSIKIDTAIKFWKLRYNGNLSKITL
SNNYELKATPDHCLLVLRDNQLKWIPAKDIKENDYIAMPFNYKVERKPISLLNLLKYLDITD
VLIEFDENSTIFEKIAEYIRNNIKTSTKYKYLRNRRVPLKYLIEWNFDLDEIEKEAKYIYKS
VAGTKKIPLFKLDERFWYFAGLVLGDGSIQDSKIRIAQTPLKDVKSILDETFPFLHNWISGN
QVIISNPIIAEILEKLGMRNGKLNGIIFSLPESYINALIAGYFDTDGCFSLLYDKKAKKHNL
RMVLTSKRRDVLEKIGIYLNSIGILNTLHKSREVYSLIISNKSLETFKEKIAKYLKIRKEAF
INGYKTYKKEHEERFECDLLPVKEVFKKLTFEKGRKEILKDSKIHIENWYKEKTNNIPREKL
KTVLRYANNSEHKEFLEKIVNGDISFVRVKKVENIPYDGYVYDLSIKHNQNFISNGVISHNC

FIG._3T

KCLTGDTKVIANGQLFELRELVEKISGGKFGPTPVKGLKVIGIDEDGKLREFEVQYVYKDKT
ERLIRIRTRLGRELKVTPYHPLLVNRRNGEIKWVKAEELKPGDKLAVPRFLPIVTGEDPLAE
WLGYFLGGGYADSKENLIMFTNEDPLLRQRFMELTEKLFSDARIREITHENGTSKVYVNSKK
ALKLVNSLGNAHIPKECWRGIRSFLRAYFDCNGGVKGNAIVLATASKEMSQEIAYALAGFGI
ISRIQEYRVIISGSDNVKKFLNEIGFINRNKLEKALKLVKKDDPGHDGLEINYELISYVKDR
LRLSFFNDKRSWSYREAKEISWELMKEIYYRLDELEKLKESLSRGILIDWNEVAKRIEEVAE
ETGIRADELLEYIEGKRKLSFKDYIKIAKVLGIDVEHTIEAMRVFARKYSSYAEIGRRLGTW
NSSVKTILESNAVNVEILERIRKIELELIEEILSDEKLKEGIAYLIFLSQNELYWDEITKVE
ELRGEFIIYDLHVPGYHNFIAGNMPTVVHNT

FIG._3U

SCVTGDTKVYTPDEREVKIRDFMNYFENGLIKEVSNRIGRDTVIAAVSFNSRIVGHPVYRLT
LESGRIIEATGDHMFLTPEGWKQTYDIKEGSEVLVKPTLEGTPYEPDPRVIIDIKEFYNFLE
KIEREHNLKPLKEAKTFRELITKDKEKILRRALELRAEIENGLTKREAEILELISADTWIPR
AELEKKARISRTRLNQILQRLEKKGYIERRIEGRKQFVRKIRNGKILRNAMDIKRILEEEFG
IKISYTTVKKLLSGNVDGMAYRILKEVKEKWLVRYDDEKAGILARVVGFILGDGHLARNGRI
WFNSSKEELEMLANDLRKLGLKPSEIIERDSSSEIQGRKVKGRIYMLYVDNAAFHALLRFWK
VEVGNKTKKGYTVPEWIKKGNLFVKREFLRGLFGADGTKPCGKRYNFNGIKLEIRAKKESLE
RTVEFLNDVADLLREFDVDSKITVSPTKEGFIIRLIVTPNDANYLNFLTRVGYAYAKDTYAR
LVGEYIRIKLAYKNIILPGIAEKAIELATVTNSTYAAKVLGVSRDFVVNRLKGTQIGITRDF
MTFEEFMKERVLNGYVIEKVIKKEKLGYLDVYDVTCARDHSFISNGLVSHNC

FIG._3V

NCLTSNSKILTDDGYYIKLEKLKEKLDLHIKIYNTEEGEKSSNILFVSERYADEKIIRIKTE
SGRVLEGSKDHPVLTLNGYVPMGMLKEGDDVIVYPYEGVEYEEPSDEIILDEDDFAEYDKQI
IKYLKDRGLLPLRMDNKNIGIIARLLGFAFGDGSIVKENGDRERLYVAFYGKRETLIKIRED
LEKLGIKASRIYSRKREVEIRNAYGDEYTSLCEDNSIKITSKAFALFMHKLGMPIGKKTEQI
YKIPEWIKKAPKWVKRNFLAGLFGADGSRAVFKNYTPLPINLTMSKSEELKENILEFLNEIK
LLLAEFDIESMIYEIKSLDGRVSYRLAIVGEESIKNFLGRINYEYSGEKKVIGLLAYEYLRR
KDIAKEIRKKCIKRAKELYKKGVTVSEMLKMDEFRNEFISKRLIERAVYENLDEDDVRISTK
FPKFEEFIEKYGVIGGFVIDKIKEIEEISYDSKLYDVGIVSKEHNFIANSIVVHNC

FIG._3W

KCVDGDTLVLTKEFGLIKIKELYEKLDGKGRKIVEGNEEWTELEKPITVYGYKDGKIVEIKA
THVYKGVSSGMVEIRTRTGRKIKVTPIHRLFTGRVTKDGLILKEVMAMHVKPGDRIAVVKKI
DGGEYIKLDSSNVGEIKVPEILNEELAEFLGYLMANGTLKSGIIEIYCDDESLLERVNSLSL
KLFGVGGRIVQKVDGKALVIQSKPLVDVLRRLGVPEDKKVENWKVPRELLLSPSNVVRAFVN
AYIKGKEEVEITLASEEGAYELSYLFAKLGIYVTISKSGEYYKVRVSRRGNLDTIPVEVNGM
PKVLPYEDFRKFAKSIGLEEVAENHLQHIIFDEVIDVRYIPEPQEVYDVTTETHNFVGGNMP
TLLHNT

FIG._3X

Intein B

MESGSPEIEKLSQSDIYWDSIVSITETGVEEVFDLTVPGPHNFVAND

Cyclic Insert (With Flag Epitope)

IIVHNSIEQGQGGGMSMDYKDDDDKMRMLEGQAGGLITSGCIS

GDSLISLASTGKRVSIKDLLDEKDFEIWAINEQTMKLESKVSRVFCT

Intein A

GKKLVYILKTRLGRTIKATANHRFLTIDGWKRLDELSLKEHIALPRK

LESSSLQLSIHGYH

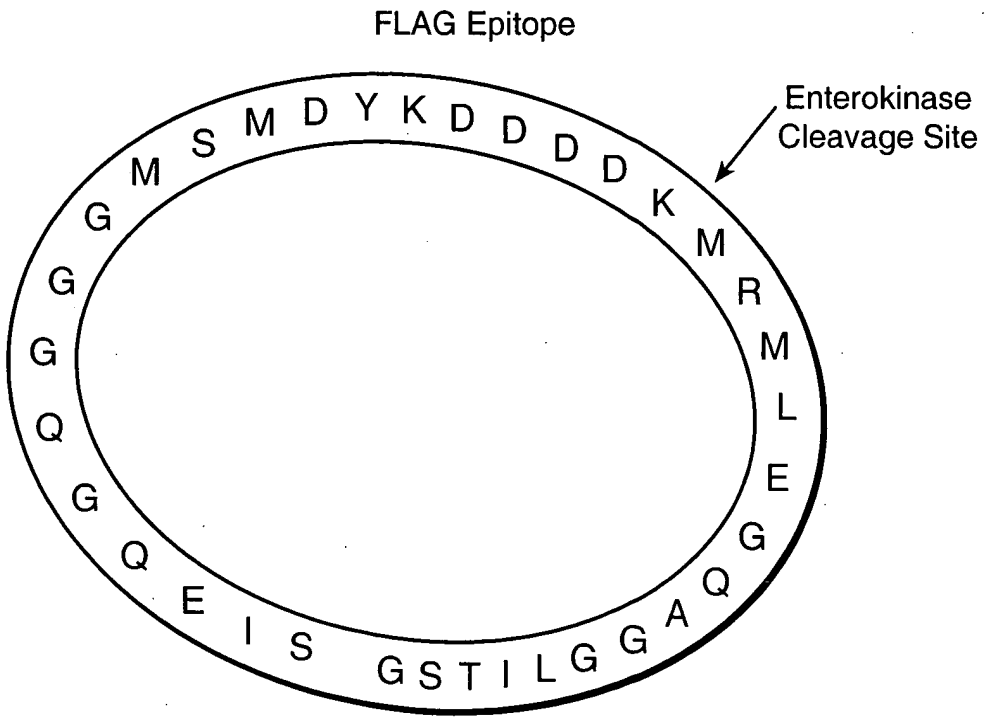


FIG._4A

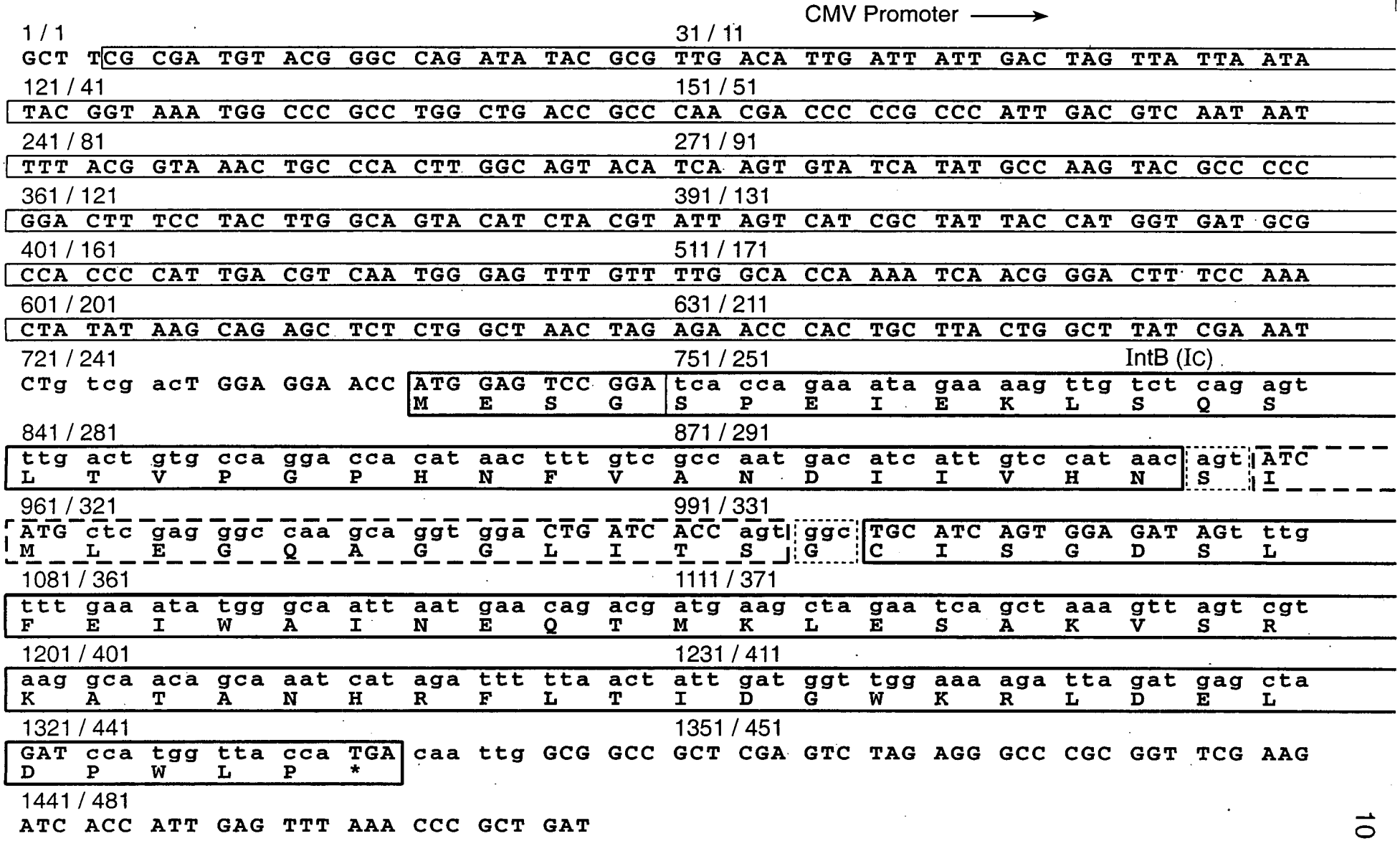

CMV Promoter
1 / 1
31 / 11
GCT TCG CGA TGT ACG GGC CAG ATA TAC GCG TTG ACA TTG ATT ATT GAC TAG TTA TTA ATA
121 / 41
151 / 51
TAC GGT AAA TGG CCC GCC TGG CTG ACC GCC CAA CGA CCC CCG CCC ATT GAC GTC AAT AAT
241 / 81
271 / 91
TTT ACG GTA AAC TGC CCA CTT GGC AGT ACA TCA AGT GTA TCA TAT GCC AAG TAC GCC CCC
361 / 121
391 / 131
GGA CTT TCC TAC TTG GCA GTA CAT CTA CGT ATT AGT CAT CGC TAT TAC CAT GGT GAT GCG
401 / 161
511 / 171
CCA CCC CAT TGA CGT CAA TGG GAG TTT GTT TTG GCA CCA AAA TCA ACG GGA CTT TCC AAA
601 / 201
631 / 211
CTA TAT AAG CAG AGC TCT CTG GCT AAC TAG AGA ACC CAC TGC TTA CTG GCT TAT CGA AAT
721 / 241
751 / 251
IntB (Ic)
CTg tcg acT GGA GGA ACC ATG GAG TCC GGA tca cca gaa ata gaa aag ttg tct cag agt
M E S G S P E I E K L S Q S
841 / 281
871 / 291
ttg act gtg cca gga cca cat aac ttt gtc gcc aat gac atc att gtc cat aac agt ATC
L T V P G P H N F V A N D I I V H N S I
961 / 321
991 / 331
ATG ctc gag ggc caa gca ggt gga CTG ATC ACC agt ggc TGC ATC AGT GGA GAT AGt ttg
M L E G Q A G G L I T S G C I S G D S L
1081 / 361
1111 / 371
ttt gaa ata tgg gca att aat gaa cag acg atg aag cta gaa tca gct aaa gtt agt cgt
F E I W A I N E Q T M K L E S A K V S R
1201 / 401
1231 / 411
aag gca aca gca aat cat aga ttt tta act att gat ggt tgg aaa aga tta gat gag cta
K A T A N H R F L T I D G W K R L D E L
1321 / 441
1351 / 451
GAT cca tgg tta cca TGA caa ttg GCG GCC GCT CGA GTC TAG AGG GCC CGC GGT TCG AAG
D P W L P *
1441 / 481
ATC ACC ATT GAG TTT AAA CCC GCT GAT

FIG._4B-1

61 / 21 91 / 31

GTA ATC AAT TAC GGG GTC ATT AGT TCA TAG CCC ATA TAT GGA GTT CCG CGT TAC ATA ACT

181 / 61 211 / 71

GAC GTA TGT TCC CAT AGT AAC GCC AAT AGG GAC TTT CCA TTG ACG TCA ATG GGT GGA CTA

301 / 101 331 / 111

TAT TGA CGT CAA TGA CGG TAA ATG GCC CGC CTG GCA TTA TGC CCA GTA CAT GAC CTT ATC

421 / 141 451 / 151

GTT TTG GCA GTA CAT CAA TGG GCG TGG ATA GCG GTT TGA CTC ACG GGG ATT TCC AAG TCT

541 / 181 571 / 191

ATG TCG TAA CAA CTC CGC CCC ATT GAC GCA AAT GGG CGG TAG GCG TGT ACG GTG GGA GGT

661 / 221 691 / 231

TAA TAC GAC TCA CTA TAG GGA GAC CCA AGC TGG CTA GTT AAG CTT cct ata cta ggA GAT

781 / 261 811 / 271 IntB (Ic)

gat	att	tac	tgg	gac	tcc	atc	gtt	tct	att	acg	gag	act	gga	gtc	gaa	gag	gtt	ttt	gat
D	I	Y	W	D	S	I	V	S	I	T	E	T	G	V	E	E	V	F	D

901 / 301 Flag Epitope Insert 931 / 311

GAA	CAA	ggc	cag	ggc	ggt	ggc	ATG	TCA	ATG	gac	tat	aaa	gat	gac	gac	gat	aag	ATG	AGG
E	Q	G	Q	G	G	G	M	S	M	D	Y	K	D	D	D	D	K	M	R

1021 / 341 1051 / 351

atc	agc	ttg	gcg	agc	aca	gga	aaa	aga	gtt	tct	att	aaa	gat	ttg	tta	gat	gaa	aaa	gat
I	S	L	A	S	T	G	K	R	V	S	I	K	D	L	L	D	E	K	D

1141 / 381 IntA (In) 1171 / 391

gta	ttt	tgt	act	ggc	aaa	aag	cta	gtt	tat	att	tta	aaa	act	cga	cta	ggt	aga	act	atc
V	F	C	T	G	K	K	L	V	Y	I	L	K	T	R	L	G	R	T	I

1261 / 421 1291 / 431

tct	tta	aaa	gag	cat	att	gct	cta	ccc	cgt	aaa	cta	gaa	agc	tcc	tct	tta	caa	tta	ATC
S	L	K	E	H	I	A	L	P	R	K	L	E	S	S	S	L	Q	L	I

1381 / 461 1411 / 471

GTA AGC CTA TCC CTA ACC CTC TCC TCG GTC TCG ATT CTA CGC GTA CCG GTC ATC ATC ACC

FIG._4B-2

FIG._4B FIG._4B-1 FIG._4B-2

ATGGAGTCCGGATCACCAGAAATAGAAAAGTTGTCTCAGAGTGATATTTACTGGGACTCCAT
CGTTTCTATTACGGAGACTGGAGTCGAAGAGGTTTTTGATTTGACTGTGCCAGGGCCCCATA
ACTTTGTGGCCAATGACATCATTGTCCATAACAGTGAGGAGGACCTGGGATCCAGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAA
CCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
TCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGG
TCGAACGGGGAATTCTCGCAGGTAGACAAGTCGATGGTGAGCAAGGGCGAGGAGCTGTTCAC
CGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC
GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTT
CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCT
ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTG
AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA
CGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGG
CCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACCTC
GAGCAAAAGCTGATATGCATCTCCGGAaATAGTTTGATCAGCTTGGCGAGCACAGGAAAAAG
AGTTTCTATTAAAGATTTGTTAGATGAAAAAGATTTTGAAATATGGGCAATTAATGAACAGA
CGATGAAGCTAGAATCAGCTAAAGTTAGTCGTGTATTTTGTACTGGCAAAAAGCTAGTTTAT
ATTTTAAAAACTCGACTAGGTAGAACTATCAAGGCAACAGCAAATCATAGATTTTTAACTAT
TGATGGTTGGAAAAGATTAGATGAGCTATCTTTAAAAGAGCATATTGCTCTACCCCGTAAAC
TAGAAAGCTCCTCTTTACAATTAGGCCTCCGCGGCCAGTACCCCTACGACGTCCCGGACTAC
GCTATCGATTAA

FIG._5A

MESGSPEIEKLSQSDIYWDSIVSITETGVEEVFDLTVPGPHNFVANDIIVHNSEEDLGSSVQ
LADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKG
SNGEFSQVDKSMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT
GKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEV
KFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDL
EQKLICISGNSLISLASTGKRVSIKDLLDEKDFEIWAINEQTMKLESAKVSRVFCTGKKLVY
ILKTRLGRTIKATANHRFLTIDGWKRLDELSLKEHIALPRKLESSSLQLGLRGQYPYDVPDY
AIDZ

FIG._5B

ATGGAGTCCGGATCACCAGAAATAGAAAAGTTGTCTCAGAGTGATATTTACTGGGACTCCAT
CGTTTCTATTACGGAGACTGGAGTCGAAGAGGTTTTTGATTTGACTGTGCCAGGGCCCCATA
ACTTTGTGGCCAATGACATCATTGTCCATAACAGTGAGGAGGACCTGGGATCCAGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAA
CCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
TCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGG
TCGAACGGGGAATTCTCGCAGGTAGACAAGTCGATGGTGAGCAAGGGCGAGGAGCTGTTCAC
CGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC
GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTT
CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCT
ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTG
AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA
TGGAGTCCGGATCACCAGAAATAGAAAAGTTGTCTCAGAGTGATATTTACTGGGACTCCATC
GTTTCTATTACGGAGACTGGAGTCGAAGAGGTTTTTGATTTGACTGTGCCAGGGCCCCATAA
CTTTGTGGCCAATGACATCATTGTCCATAACAGTGAGGAGGACCTGGGATCCAGCGTGCAGC
TCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAAC
CACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGT
CCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGGT
CGAACGGGGAATTCTCGCAGGTAGACAAGTCGATGGTGAGCAAGGGCGAGGAGCTGTTCACC
GGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTC
CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCG
GCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTC
AGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTA
CGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGA
AGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGAC
GGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGC
CGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACCTCG
AGCAAAAGCTGATATGCATCTCCGGAAATAGTTTGATCAGCTTGGCGAGCACAGGAAAAAGA
GTTTCTATTAAAGATTTGTTAGATGAAAAAGATTTTGAAATATGGGCAATTAATGAACAGAC
GATGAAGCTAGAATCAGCTAAAGTTAGTCGTGTATTTTGTACTGGCAAAAAGCTAGTTTATA
TTTTAAAAACTCGACTAGGTAGAACTATCAAGGCAACAGCAAATCATAaATTTTTAACTATT
GATGGTTGGAAAAGATTAGATGAGCTATCTTTAAAAGAGCATATTGCTCTACCCCGTAAACT
AGAAAGCTCCTCTTTACAATTAGGCCTCCGCGGCCAGTACCCCTACGACGTCCCGGACTACG
CTATCGATTAA

FIG._5C

MESGSPEIEKLSQSDIYWDSIVSITETGVEEVFDLTVPGPHNFVANDIIVHNSEEDLGSSVQ
LADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKG
SNGEFSQVDKSMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT
GKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEV
KFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDL
EQKLICISGNSLISLASTGKRVSIKDLLDEKDFEIWAINEQTMKLESAKVSRVFCTGKKLVY
ILKTRLGRTIKATANHKFLTIDGWKRLDELSLKEHIALPRKLESSSLQLGLRGQYPYDVPDY
AIDZ

FIG._5D

ATGGAGTCCGGATCACCAGAAATAGAAAAGTTGTCTCAGAGTGATATTTACTGGGACTCCAT
CGTTTCTATTACGGAGACTGGAGTCGAAGAGGTTTTTGATTTGACTGTGCCAGGGCCCCATA
ACTTTGTGGCCAATGACATCATTGTCCATAACAGTGAGGAGGACCTGGGATCCAGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAA
CCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
TCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGG
TCGAACGGGGAATTCTCGCAGGTAGACAAGTCGATGGTGAGCAAGGGCGAGGAGCTGTTCAC
CGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC
GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTT
CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCT
ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTG
AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA
CGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGG
CCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACCTC
GAGCAAAAGCTGATATGCATCTCCGGAaATAGTTTGATCAGCTTGGCGAGCACAGGAAAAAG
AGTTTCTATTAAAGATTTGTTAGATGAAAAAGATTTTGAAATATGGGCAgTTAATGAACAGA
CGATGAAGCTAGAATCAGCTAAAGTTAGTCGTGTATTTTGTACTGGCAAAAAGCTAGTTTAT
ATTTTAAAAACTCGACTAGGTAGAACTATCAAGGCAACAGCAAATCATAGATTTTTAACTAT
TGATGGTTGGAAAAGATTAGATGAGCTATCTTTAAAAGAGCATATTGCTCTACCCCGTAAAC
TAGAAAGCTCCTCTTTACAATTAGGCCTCCGCGGCCAGTACCCCTACGACGTCCCGGACTAC
GCTATCGATTAA

FIG._5E

MESGSPEIEKLSQSDIYWDSIVSITETGVEEVFDLTVPGPHNFVANDIIVHNSEEDLGSSVQ
LADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKG
SNGEFSQVDKSMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT
GKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEV
KFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDL
EQKLICISGNSLISLASTGKRVSIKDLLDEKDFEIWAVNEQTMKLESAKVSRVFCTGKKLVY
ILKTRLGRTIKATANHRFLTIDGWKRLDELSLKEHIALPRKLESSSLQLGLRGQYPYDVPDY
AIDZ

FIG._5F

ATGGAGTCCGGATCACCAGAAATAGAAAAGTTGTCTCAGAGTGATATTTACTGGGACTCCAT
CGTTTCTATTACGGAGACTGGAGTCGAAGAGGTTTTTGATTTGGcCcGTGCCAGGGCCCCATA
ACTTTGTGGCCAATGACATCATTGTCCATAACAGTGAGGAGGACCTGGGATCCAGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAA
CCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
TCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGG
TCGAACGGGGAATTCTCGCAGGTAGACAAGTCGATGGTGAGCAAGGGCGAGGAGCTGTTCAC
CGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC
GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTT
CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCT
ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTG
AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA
CGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGG
CCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACCTC
GAGCAAAAGCTGATATGCATCTCCGGAAATAGTTTGATCAGCTTGGCGAGCACAGGAAAAAG
AGTTTCTATTAAAGATTTGTTAGATGAAAAAGATTTTGAAATATGGGCAATTAATGAACAGA
CGATGAAGCTAGAATCAGCTAAAGTTAGTCGTGTATTTTGTACTGGCAAAAAGCTAGTTTAT
ATTTTAAAAACTCGACTAGGTAGAACTATCAAGGCAACAGCAAATCATAGATTTTTAACTAT
TGATGGTTGGAAAAGATTAGATGAGCTATCTTTAAAAGAGCATATTGCTCTACCCCGTAAAC
TAGAAAGCTCCTCTTTACAATTAGGCCTCCGCGGCCAGTACCCCTACGACGTCCCGGACTAC
GCTATCGATTAA

FIG._5G

MESGSPEIEKLSQSDIYWDSIVSITETGVEEVFDLAVPGPHNFVANDIIVHNSEEDLGSSVQ
LADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKG
SNGEFSQVDKSMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT
GKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEV
KFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDL
EQKLICISGNSLISLASTGKRVSIKDLLDEKDFEIWAINEQTMKLESAKVSRVFCTGKKLVY
ILKTRLGRTIKATANHRFLTIDGWKRLDELSLKEHIALPRKLESSSLQLGLRGQYPYDVPDY
AIDZ

FIG._5H

ATGGAGTCCGGATCACCAGAAATAGAAAAGTTGTCTCAGAGTGATATTTACTGGGACTCCAT
CGTTcCTATTACGGAGACTGGAGTCGAAGAGGTTTTTGATTTGACTGTGCCAGGGCCCCATA
ACTTTGTGGCCAATGACATCATTGTCCATAACAGTGAGGAGGACCTGGGATCCAGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAA
CCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
TCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGG
TCGAACGGGGAATTCTCGCAGGTAGACAAGTCGATGGTGAGCAAGGGCGAGGAGCTGTTCAC
CGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC
GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTT
CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCT
ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTG
AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA
CGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGG
CCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACCTC
GAGCAAAAGCTGATATGCATCTCCGGAaATAGTTTGATCAGCTTGGCGAGCACAGGAAAAAG
AGTTTCTATTAAAGATTTGTTAGATGAAAAAGATTTTGAAATATGGGCAATTAATGAACAGA
CGATGAAGCTAGAATCAGCTAAAGTTAGTCGTGTATTTTGTACTGGCAAAAAGCTAGTTTAT
ATTTTAAAAACTCGACTAGGTAGAACTATCAAGGCAACAGCAAATCATAGATTTTTAACTAT
TGATGGTTGGAAAAGATTAGATGAGCTATCTTTAAAAGAGCATATTGCTCTACCCCGTAAAC
TAGAAAGCTCCTCTTTACAATTAGGCCTCCGCGGCCAGTACCCCTACGACGTCCCGGACTAC
GCTATCGATTAA

FIG._5I

MESGSPEIEKLSQSDIYWDSIVPITETGVEEVFDLTVPGPHNFVANDIIVHNSEEDLGSSVQ
LADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKG
SNGEFSQVDKSMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT
GKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEV
KFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDL
EQKLICISGNSLISLASTGKRVSIKDLLDEKDFEIWAINEQTMKLESAKVSRVFCTGKKLVY
ILKTRLGRTIKATANHRFLTIDGWKRLDELSLKEHIALPRKLESSSLQLGLRGQYPYDVPDY
AIDZ

FIG._5J

ATGGAGTCCGGATCACCAGAAATAGAAAAGTTGTCTCAGAGTGATATTTACTGGGACTCCAT
CGTTTCTATTACGGAGACTGGAGTCGAAGAGGTTTTTGATTTGACTGTGCCAGGGCCCCATA
ACTTTGTGGCCAATGACATCATTGTCCATAACAGTGAGGAGGACCTGGGATCCAGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAA
CCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
TCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGG
TCGAACGGGGAATTCTCGCAGGTAGACAAGTCGATGGTGAGCAAGGGCGAGGAGCTGTTCAC
CGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC
GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTT
CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCT
ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTG
AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA
CGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGG
CCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACCTC
GAGCAAAAGCTGATATGCATCTCCGGAaATAGTTTGATCAGCTTGGCGAGCACAGGAAAAAG
AGTTTCTATTAAAGATTTGTTAGATGAAAAAGATTTTGAAATATGGGCAATTAATGAACAGA
CGATGAAGCTAGAATCAGCTAAAGTTAGTCGTGTATTTTGTACTGGCAAAAgGCTAGTTTAT
ATTTTAAAAACTCGACTAGGTAGAACTATCAAGGCAACAGCAAATCATAGATTTTTAACTAT
TGATGGTTGGAAAAGATTAGATGAGCTATCTTTAAAAGAGCATATTGCTCTACCCCGTAAAC
TAGAAAGCTCCTCTTTACAATTAGGCCTCCGCGGCCAGTACCCCTACGACGTCCCGGACTAC
GCTATCGATTAA

FIG._5K

MESGSPEIEKLSQSDIYWDSIVSITETGVEEVFDLTVPGPHNFVANDIIVHNSEEDLGSSVQ
LADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKG
SNGEFSQVDKSMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT
GKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEV
KFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDL
EQKLICISGNSLISLASTGKRVSIKDLLDEKDFEIWAINEQTMKLESAKVSRVFCTGKRLVY
ILKTRLGRTIKATANHRFLTIDGWKRLDELSLKEHIALPRKLESSSLQLGLRGQYPYDVPDY
AIDZ

FIG._5L

ATGGAGTCCGGATCACCAGAAATAGAAAAGTTGTCTCAGAGTGATATTTACTGGGACTCCAT
CGTTTCTATTACGGAGACTGGAGTCGAAGAGGTTTTTGATTTGACTGTGCCAGGGCCCCATA
ACTTTGTGGCCAATGACATCATTGTCCATAACAGTGAGGAGGACCTGGGATCCAGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAA
CCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
TCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGG
TCGAACGGGGAATTCTCGCAGGTAGACAAGTCGATGGTGAGCAAGGGCGAGGAGCTGTTCAC
CGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC
GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTT
CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCT
ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTG
AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA
CGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGG
CCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACCTC
GAGCAAAAGCTGATATGCATCTCCGGAGATAGTTTGATCAGCTTGGCGAGCACAGGAAAAAG
AGTTTCTATTAAAGATTTGTTAGATGAAAAAGATTTTGAAATATGGGCAATTAATGAACAGA
CGATGAAGCTAGAATCAGCTAAAGTTAGTCGTGTATTTTGTACTGGCAAAAAGCTAGTTTAT
ATTTTAAAAACTCGACTAGGTAGAACTATCAAGGCAACAGCAAATCATAaATTTTTAACTAT
TGATGGTTGGAAAAGATTAGATGAGCTATCTTTAAAAGAGCATATTGCTCTACCCCGTAAAC
TAGAAAGCTCCTCTTTACAATTAGGCCTCCGCGGCCAGTACCCCTACGACGTCCCGGACTAC
GCTATCGATTAA

FIG._5M

MESGSPEIEKLSQSDIYWDSIVSITETGVEEVFDLTVPGPHNFVANDIIVHNSEEDLGSSVQ
LADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKG
SNGEFSQVDKSMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT
GKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEV
KFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDL
EQKLICISGDSLISLASTGKRVSIKDLLDEKDFEIWAINEQTMKLESAKVSRVFCTGKKLVY
ILKTRLGRTIKATANHKFLTIDGWKRLDELSLKEHIALPRKLESSSLQLGLRGQYPYDVPDY
AIDZ

FIG._5N

ATGGAGTCCGGATCACCAGAAATAGAAAAGTTGTCTCAGAGTGATATTTACTGGGACTCCAT
CGTTcCTATTACGGAGACTGGAGTCGAAGAGGTTTTTGATTTGACTGTGCCAGGGCCCCATA
ACTTTGTGGCCAATGACATCATTGTCCATAACAGTGAGGAGGACCTGGGATCCAGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAA
CCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
TCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGG
TCGAACGGGGAATTCTCGCAGGTAGACAAGTCGATGGTGAGCAAGGGCGAGGAGCTGTTCAC
CGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC
GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTT
CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCT
ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTG
AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA
CGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGG
CCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACCTC
GAGCAAAAGCTGATATGCATCTCCGGAGATAGTTTGATCAGCTTGGCGAGCACAGGAAAAAG
AGTTTCTATTAAAGATTTGTTAGATGAAAAAGATTTTGAAATATGGGCAATTAATGAACAGA
CGATGAAGCTAGAATCAGCTAAAGTTAGTCGTGTATTTTGTACTGGCAAAAAGCTAGTTTAT
ATTTTAAAAACTCGACTAGGTAGAACTATCAAGGCAACAGCAAATCATAGATTTTTAACTAT
TGATGGTTGGAAAAGATTAGATGAGCTATCTTTAAAAGAGCATATTGCTCTACCCCGTAAAC
TAGAAAGCTCCTCTTTACAATTAGGCCTCCGCGGCCAGTACCCCTACGACGTCCCGGACTAC
GCTATCGATTAA

FIG._5O

MESGSPEIEKLSQSDIYWDSIVPITETGVEEVFDLTVPGPHNFVANDIIVHNSEEDLGSSVQ
LADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKG
SNGEFSQVDKSMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT
GKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEV
KFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDL
EQKLICISGDSLISLASTGKRVSIKDLLDEKDFEIWAINEQTMKLESAKVSRVFCTGKKLVY
ILKTRLGRTIKATANHRFLTIDGWKRLDELSLKEHIALPRKLESSSLQLGLRGQYPYDVPDY
AIDZ

FIG._5P

ATGGAGTCCGGATCACCAGAAATAGAAAAGTTGTCTCAGAGTGATATTTACTGGGACTCCAT
CGTTTCTATTACGGAGACTGGAGTCGAAGAGGTTTTTGATTTGACTGTGCCAGGGCCCCATA
ACTTTGTGGCCAATGACATCATTGTCCATAACAGTGAGGAGGACCTGGGATCCAGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAA
CCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
TCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGG
TCGAACGGGGAATTCTCGCAGGTAGACAAGTCGATGGTGAGCAAGGGCGAGGAGCTGTTCAC
CGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC
GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTT
CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCT
ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTG
AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA
CGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGG
CCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACCTC
GAGCAAAAGCTGATATGCATCTCCGGAGATAGTTTGATCAGCTTGGCGAGCACAGGAAAAAG
AGTTTCTATTAAAGATTTGTTAGATGAAAAAGATTTTGAAATATGGGCAGTTAATGAACAGA
CGATGAAGCTAGAATCAGCTAAAGTTAGTCGTGTATTTTGTACTGGCAAAAAGCTAGTTTAT
ATTTTAAAAACTCGACTAGGTAGAACTATCAAGGCAACAGCAAATCATAGATTTTTAACTAT
TGATGGTTGGAAAAGATTAGATGAGCTATCTTTAAAAGAGCATATTGCTCTACCCCGTAAAC
TAGAAAGCTCCTCTTTACAATTAGGCCTCCGCGGCCAGTACCCCTACGACGTCCCGGACTAC
GCTATCGATTAA

FIG._5Q

MESGSPEIEKLSQSDIYWDSIVSITETGVEEVFDLTVPGPHNFVANDIIVHNSEEDLGSSVQ
LADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKG
SNGEFSQVDKSMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT
GKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEV
KFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDL
EQKLICISGDSLISLASTGKRVSIKDLLDEKDFEIWAVNEQTMKLESAKVSRVFCTGKKLVY
ILKTRLGRTIKATANHRFLTIDGWKRLDELSLKEHIALPRKLESSSLQLGLRGQYPYDVPDY
AIDZ

FIG._5R

CMV Promoter ⟶

1 / 1 — 31 / 11

GCT TCG CGA TGT ACG GGC CAG ATA TAC GCG TTG ACA TTG ATT ATT GAC TAG TTA TTA ATA

121 / 41 — 151 / 51

TAC GGT AAA TGG CCC GCC TGG CTG ACC GCC CAA CGA CCC CCG CCC ATT GAC GTC AAT AAT

241 / 81 — 271 / 91

TTT ACG GTA AAC TGC CCA CTT GGC AGT ACA TCA AGT GTA TCA TAT GCC AAG TAC GCC CCC

361 / 121 — 391 / 131

GGA CTT TCC TAC TTG GCA GTA CAT CTA CGT ATT AGT CAT CGC TAT TAC CAT GGT GAT GCG

401 / 161 — 511 / 171

CCA CCC CAT TGA CGT CAA TGG GAG TTT GTT TTG GCA CCA AAA TCA ACG GGA CTT TCC AAA

601 / 201 — 631 / 211

CTA TAT AAG CAG AGC TCT CTG GCT AAC TAG AGA ACC CAC TGC TTA CTG GCT TAT CGA AAT

721 / 241 — 751 / 251 — IntB (IC)

CTg	tcg	acT	GGA	GGA	ACC	ATG	GAG	TCC	GGA	tca	cca	gaa	ata	gaa	aag	ttg	tct	cag	agt	
						M	E	S	G	S	P	E	I	E	K	L	S	Q	S	

841 / 281 — 871 / 291

ttg	act	gtg	cca	gga	cca	cat	aac	ttt	gtc	gcc	aat	gac	atc	att	gtc	cat	aac	agt	ATC	
L	T	V	P	G	P	H	N	F	V	A	N	D	I	I	V	H	N	S	I	

961 / 321 — 991 / 331

ATG	ctc	gag	ggc	caa	gca	ggt	gga	CTG	ATC	ACC	agt	ggc	TGC	ATC	AGT	GGA	GAT	AGt	ttg	
M	L	E	G	Q	A	G	G	L	I	T	S	G	C	I	S	G	D	S	L	

1081 / 361 — 1111 / 371

ttt	gaa	ata	tgg	gca	att	aat	gaa	cag	acg	atg	aag	cta	gaa	tca	gct	aaa	gtt	agt	cgt	
F	E	I	W	A	I	N	E	Q	T	M	K	L	E	S	A	K	V	S	R	

1201 / 401 — 1231 / 411

aag	gca	aca	gca	aat	cat	aga	ttt	tta	act	att	gat	ggt	tgg	aaa	aga	tta	gat	gag	cta	
K	A	T	A	N	H	R	F	L	T	I	D	G	W	K	R	L	D	E	L	

FIG._6A

61 / 21 — 91 / 31

GTA ATC AAT TAC GGG GTC ATT AGT TCA TAG CCC ATA TAT GGA GTT CCG CGT TAC ATA ACT

181 / 61 — 211 / 71

GAC GTA TGT TCC CAT AGT AAC GCC AAT AGG GAC TTT CCA TTG ACG TCA ATG GGT GGA CTA

301 / 101 — 331 / 111

TAT TGA CGT CAA TGA CGG TAA ATG GCC CGC CTG GCA TTA TGC CCA GTA CAT GAC CTT ATC

421 / 141 — 451 / 151

GTT TTG GCA GTA CAT CAA TGG GCG TGG ATA GCG GTT TGA CTC ACG GGG ATT TCC AAG TCT

541 / 181 — 571 / 191

ATG TCG TAA CAA CTC CGC CCC ATT GAC GCA AAT GGG CGG TAG GCG TGT ACG GTG GGA GGT

661 / 221 — 691 / 231

TAA TAC GAC TCA CTA TAG GGA GAC CCA AGC TGG CTA GTT AAG CTT cct ata cta ggA GAT

781 / 261 — 811 / 271 — IntB (Ic)

gat	att	tac	tgg	gac	tcc	atc	gtt	tct	att	acg	gag	act	gga	gtc	gaa	gag	gtt	ttt	gat
D	I	Y	W	D	S	I	V	S	I	T	E	T	G	V	E	E	V	F	D

901 / 301 — Flag Epitope Insert — 931 / 311

GAA	CAA	ggc	cag	ggc	ggt	ggc	ATG	TCA	ATG	gac	tat	aaa	gat	gac	gac	gat	aag	ATG	AGG
E	Q	G	Q	G	G	G	M	S	M	D	Y	K	D	D	D	D	K	M	R

1021 / 341 — 1051 / 351

atc	agc	ttg	gcg	agc	aca	gga	aaa	aga	gtt	tct	att	aaa	gat	ttg	tta	gat	gaa	aaa	gat
I	S	L	A	S	T	G	K	R	V	S	I	K	D	L	L	D	E	K	D

1141 / 381 — IntA (In) — 1171 / 391

gta	ttt	tgt	act	ggc	aaa	aag	cta	gtt	tat	att	tta	aaa	act	cga	cta	ggt	aga	act	atc
V	F	C	T	G	K	K	L	V	Y	I	L	K	T	R	L	G	R	T	I

1261 / 421 — 1291 / 431

tct	tta	aaa	gag	cat	att	gct	cta	ccc	cgt	aaa	cta	gaa	agc	tcc	tct	tta	caa	tta	ATC
S	L	K	E	H	I	A	L	P	R	K	L	E	S	S	S	L	Q	L	I

FIG._6B

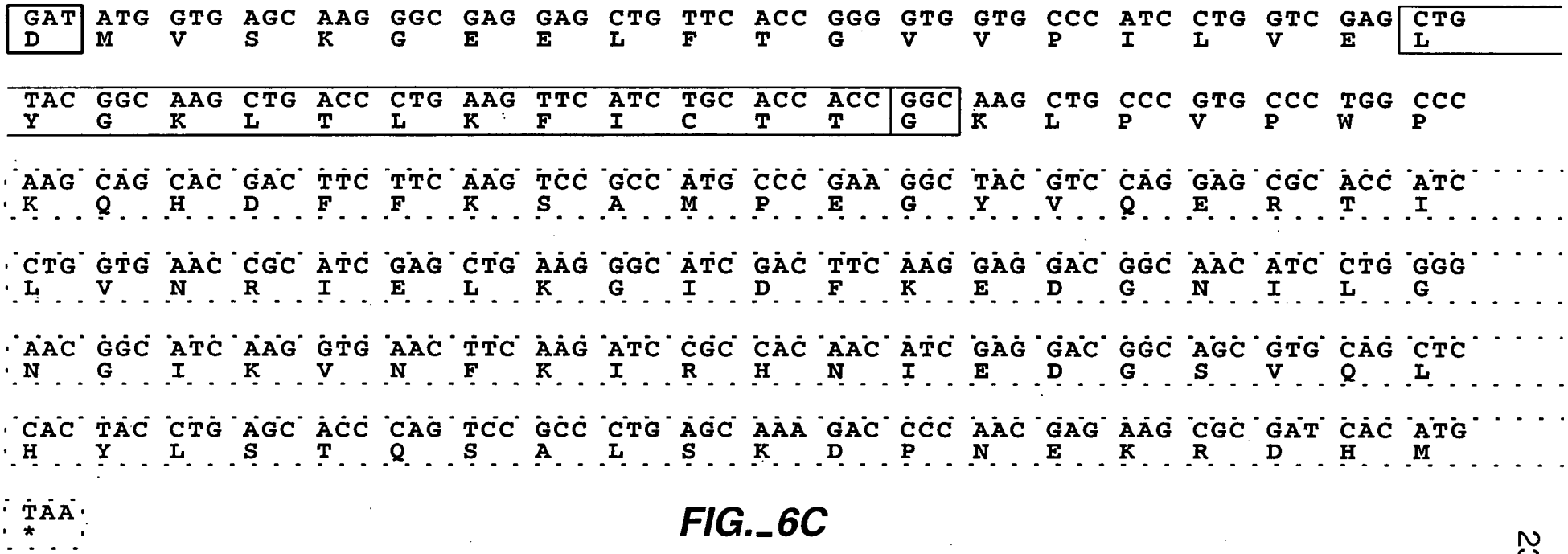

GAT ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG GTC GAG CTG
D M V S K G E E L F T G V V P I L V E L
TAC GGC AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC
Y G K L T L K F I C T T G K L P V P W P
AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG CGC ACC ATC
K Q H D F F K S A M P E G Y V Q E R T I
CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG
L V N R I E L K G I D F K E D G N I L G
AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC GTG CAG CTC
N G I K V N F K I R H N I E D G S V Q L
CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG
H Y L S T Q S A L S K D P N E K R D H M
TAA
*

FIG._6C

GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC
D G D V N G H K F S V S G E G E G D A T

GFP GFP

ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG
T L V T T L T Y G V Q C F S R Y P D H M

TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC
F F K D D G N Y K T R A E V K F E G D T

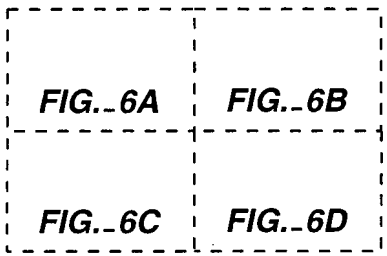
CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG
H K L E Y N Y N S H N V Y I M A D K Q K

GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC GAC AAC
A D H Y Q Q N T P I G D G P V L L P D N

GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG
V L L E F V T A A G I T L G M D E L Y K

FIG._6D

FIG._6



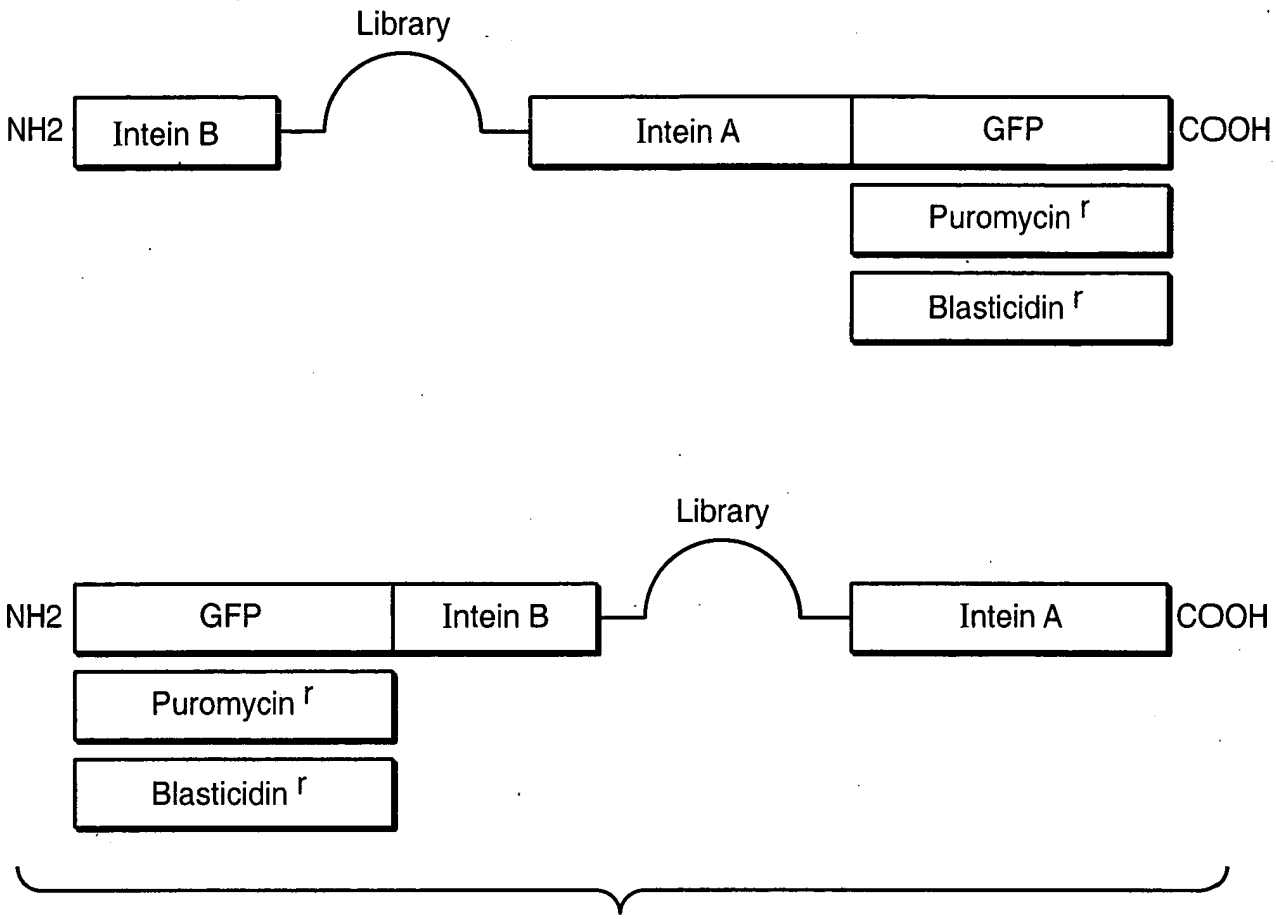

Library
NH2
Intein B
Intein A
GFP
COOH
Puromycin r
Blasticidin r
Library
NH2
GFP
Intein B
Intein A
COOH
Puromycin r
Blasticidin r

FIG._7

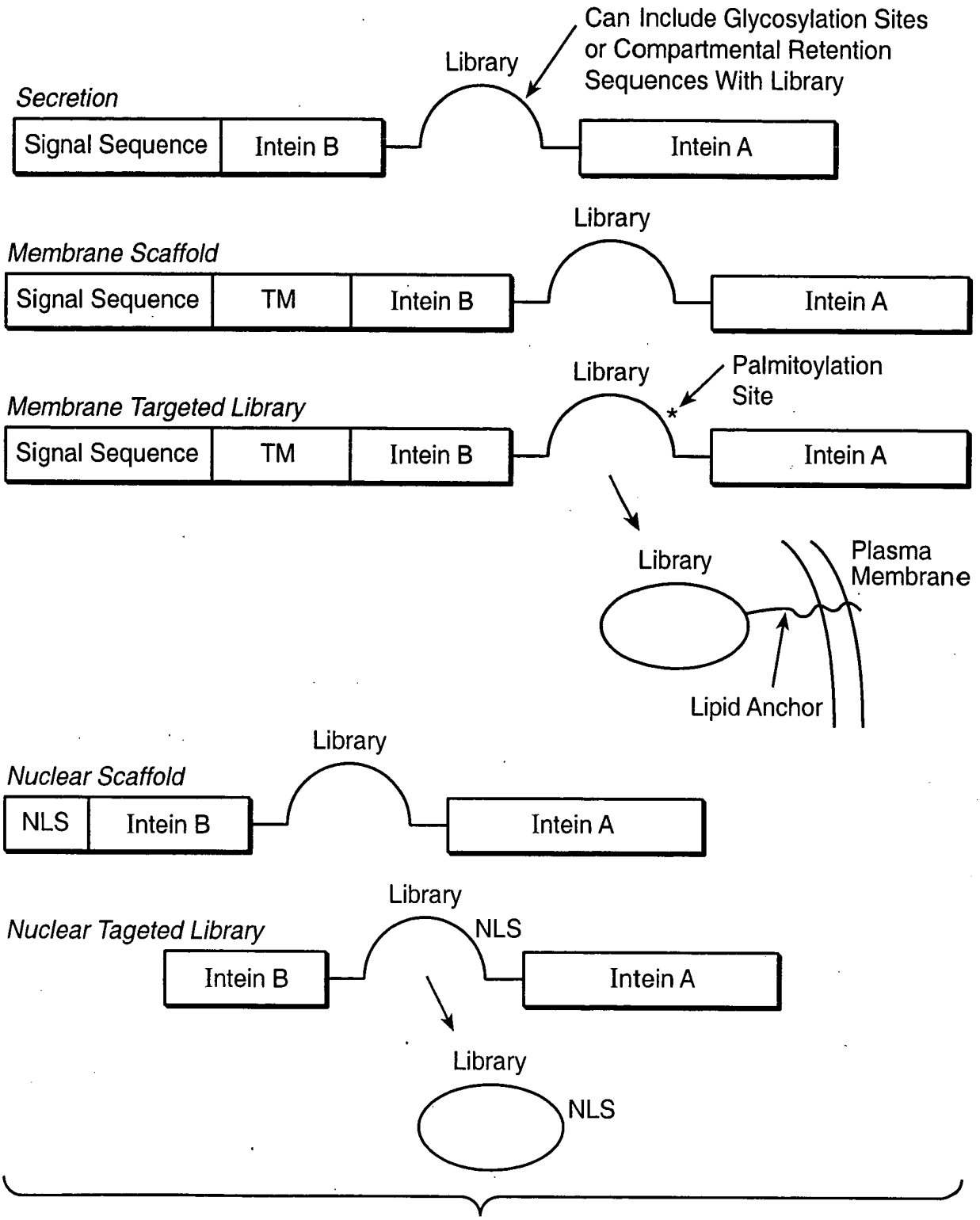

Can Include Glycosylation Sites
or Compartmental Retention
Sequences With Library
Library
Secretion
Signal Sequence
Intein B
Intein A
Library
Membrane Scaffold
Signal Sequence
TM
Intein B
Intein A
Library
Palmitoylation
Site
Membrane Targeted Library
Signal Sequence
TM
Intein B
Intein A
Library
Plasma
Membrane
Lipid Anchor
Library
Nuclear Scaffold
NLS
Intein B
Intein A
Library
Nuclear Tageted Library
NLS
Intein B
Intein A
Library
NLS

FIG._8

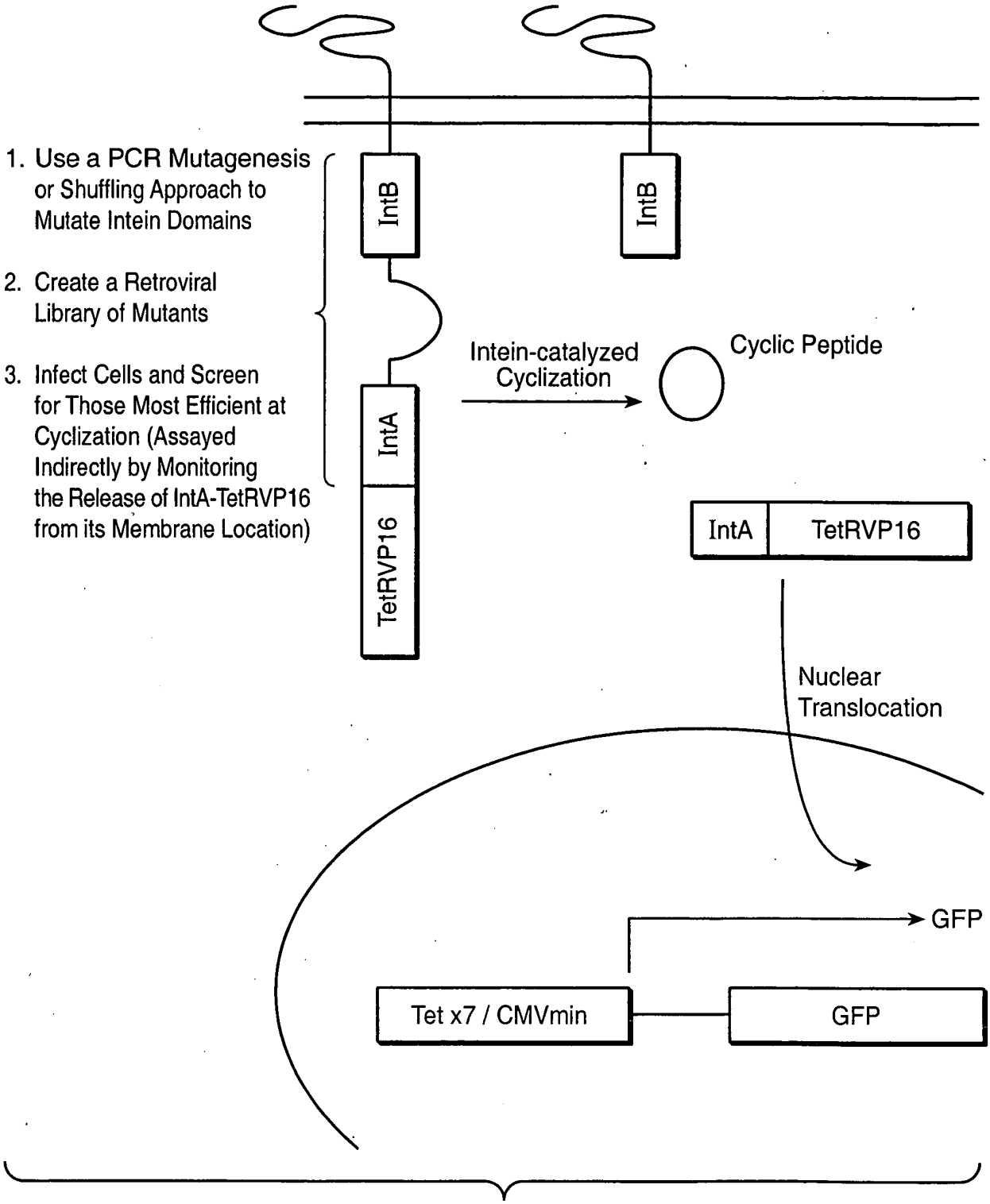

1. Use a PCR Mutagenesis or Shuffling Approach to Mutate Intein Domains
2. Create a Retroviral Library of Mutants
3. Infect Cells and Screen for Those Most Efficient at Cyclization (Assayed Indirectly by Monitoring the Release of IntA-TetRVP16 from its Membrane Location)
IntB
IntB
IntA
TetRVP16
Intein-catalyzed Cyclization
Cyclic Peptide
IntA
TetRVP16
Nuclear Translocation
GFP
Tet x7 / CMVmin
GFP

FIG._9

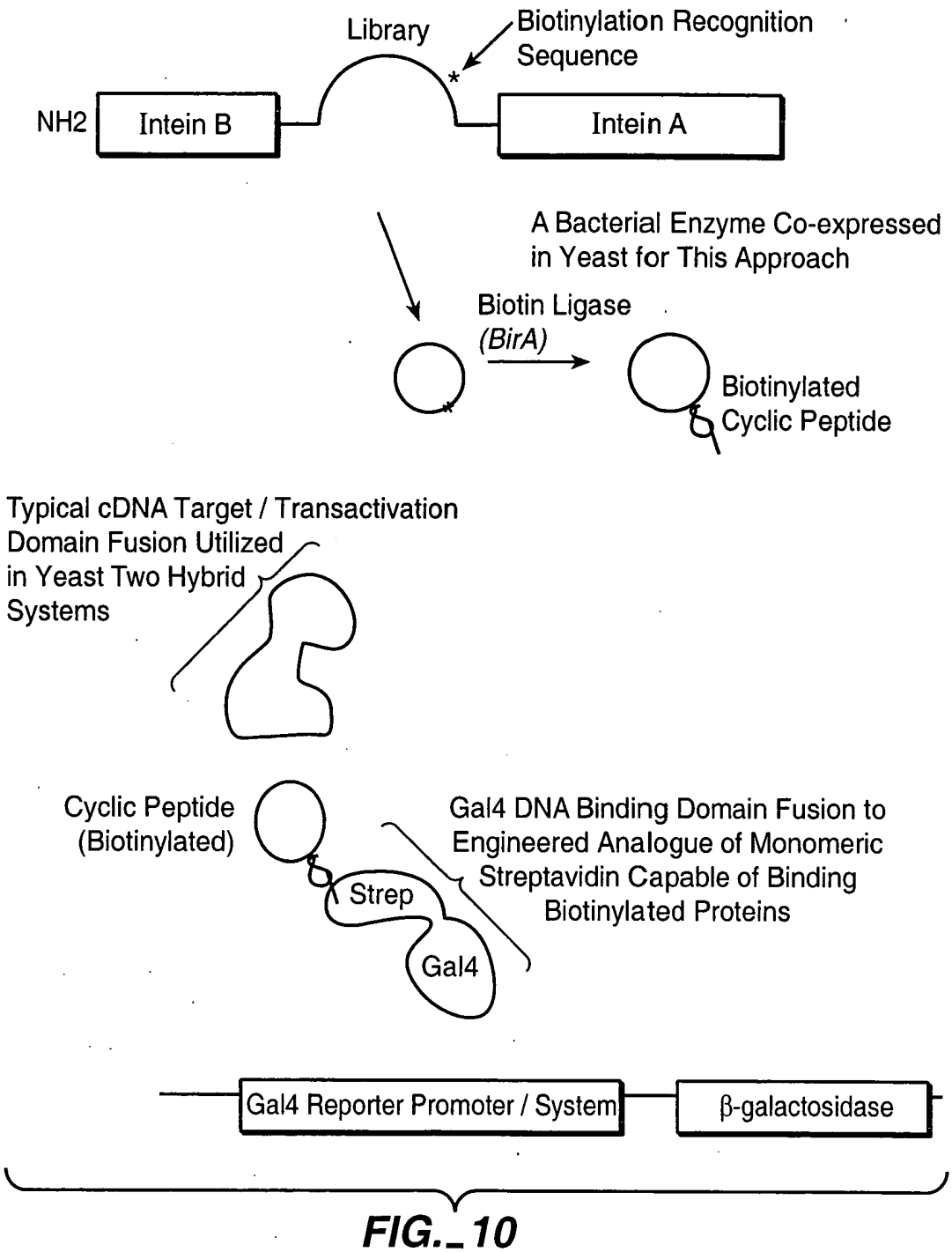

Library
Biotinylation Recognition Sequence
NH2
Intein B
Intein A
A Bacterial Enzyme Co-expressed in Yeast for This Approach
Biotin Ligase
(BirA)
Biotinylated Cyclic Peptide
Typical cDNA Target / Transactivation Domain Fusion Utilized in Yeast Two Hybrid Systems
Cyclic Peptide (Biotinylated)
Gal4 DNA Binding Domain Fusion to Engineered Analogue of Monomeric Streptavidin Capable of Binding Biotinylated Proteins
Strep
Gal4
Gal4 Reporter Promoter / System
β-galactosidase

FIG._10

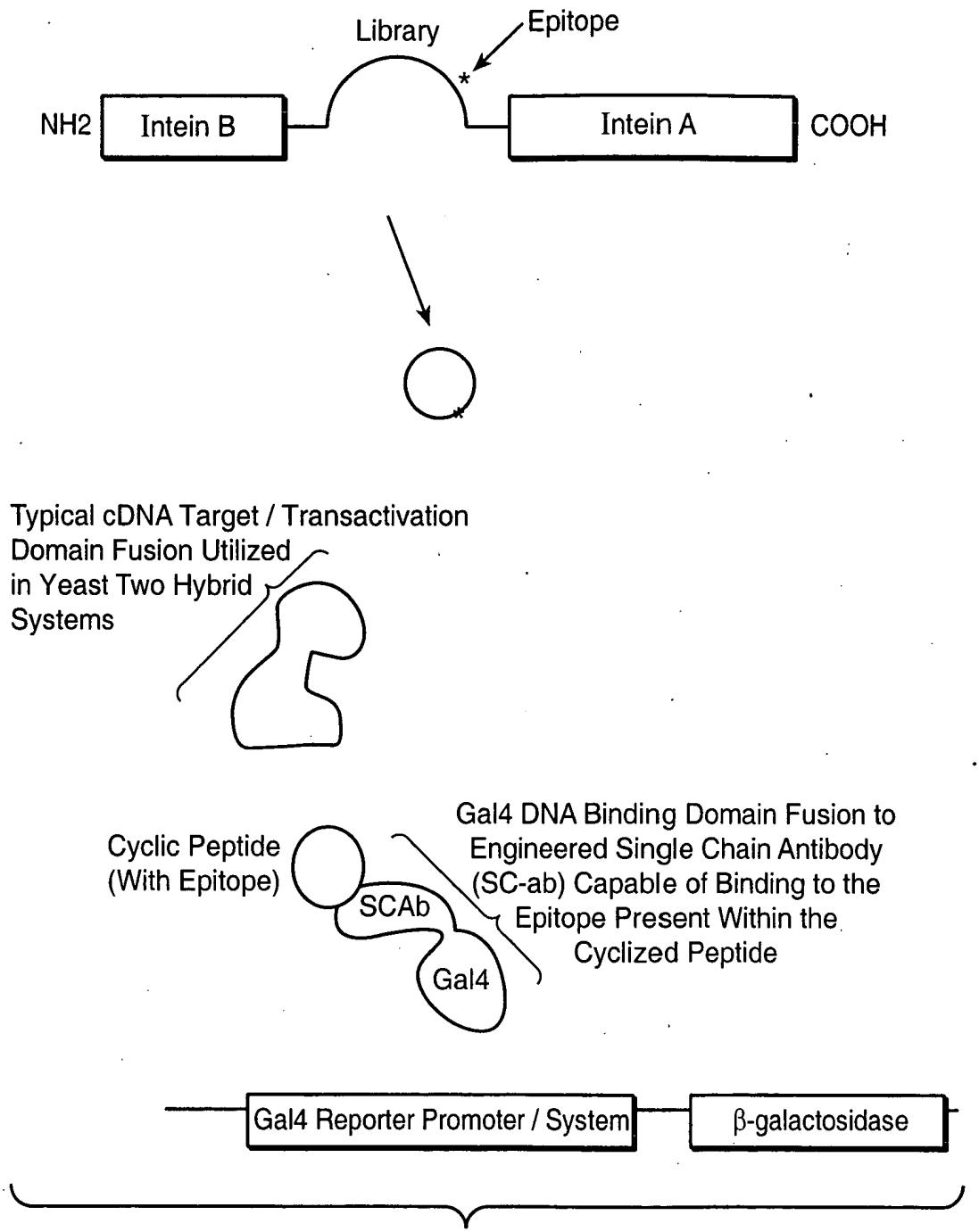

Library
Epitope
NH2
Intein B
Intein A
COOH
Typical cDNA Target / Transactivation Domain Fusion Utilized in Yeast Two Hybrid Systems
Cyclic Peptide (With Epitope)
SCAb
Gal4
Gal4 DNA Binding Domain Fusion to Engineered Single Chain Antibody (SC-ab) Capable of Binding to the Epitope Present Within the Cyclized Peptide
Gal4 Reporter Promoter / System
β-galactosidase

FIG._11

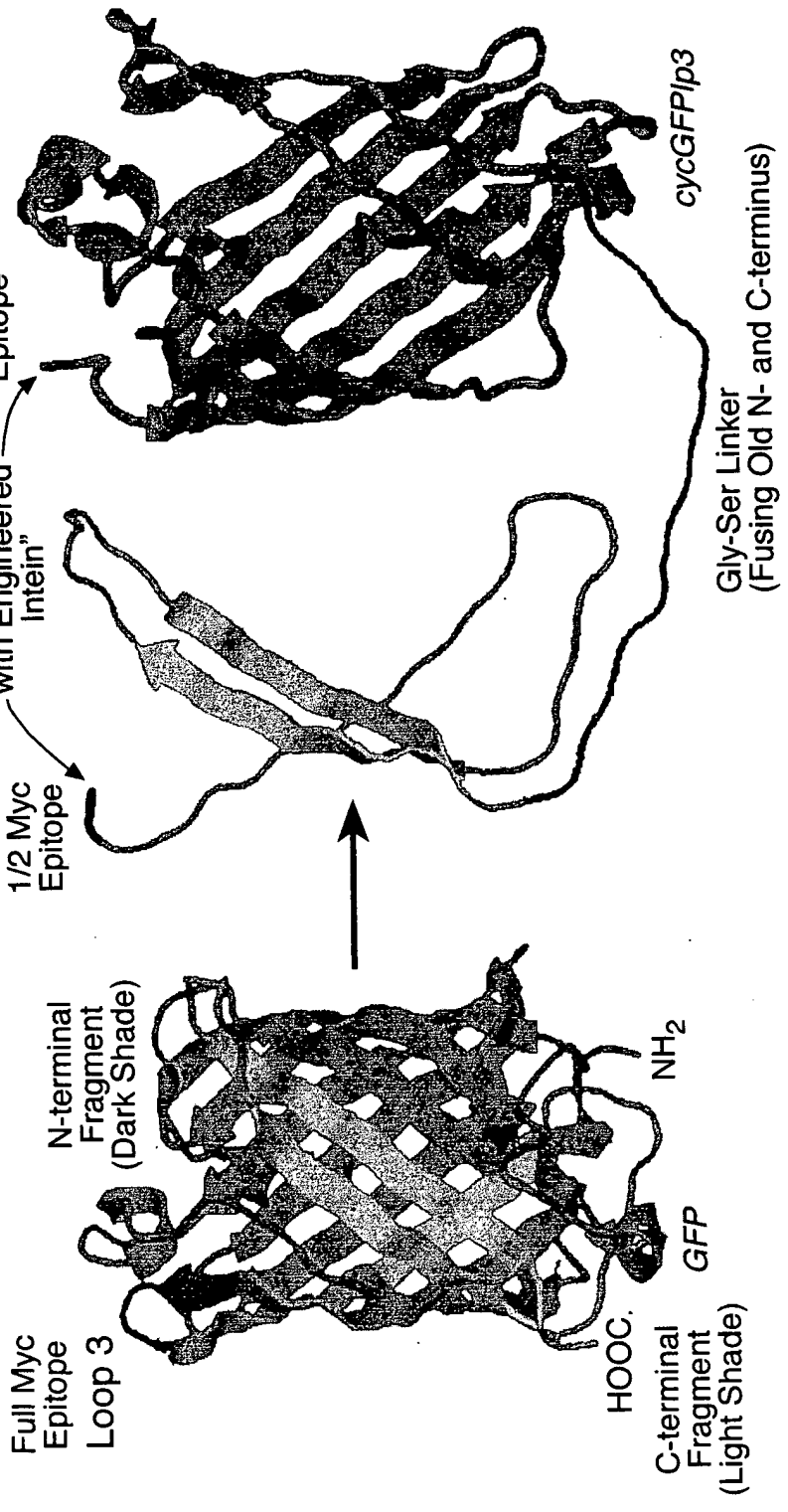

Full Myc Epitope Loop 3
N-terminal Fragment (Dark Shade)
C-terminal Fragment (Light Shade)
HOOC.
GFP
NH2
1/2 Myc Epitope
"Junction Points with Engineered Intein"
1/2 Myc Epitope
Gly-Ser Linker (Fusing Old N- and C-terminus)
cycGFPip3

FIG._12A

IntB (Ic)

MESGSPEIEKLSQSDIYWDSIVSITETGVEEVFDLTVPGP

myc$^{6-10}$

HNFVANDIIVHNSEEDLGSSVQLADHYQQNTPIGDGPVLL

PDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDE

Gly-Ser Linker

LYKGSNGEFSQVDKSMVSKGEELFTGVVPILVELDGDVNG

GFP$^{6-1-173}$

HKFSVSGEGEGDATYGLKTLKFICTTGKLPVPWPTLVTTL

TYGLQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDG

NYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYN

myc$^{1-5}$

YNSHNVYIMADKQKNGIKVNFKIRHNIEDLEQKLICISGD

SLISLASTGKRVSIKDLLDEKDFEIWAINEQTMKLESAKV

IntA (IN)

SRVFCTGKKLVYILKTRLGRTIKATANHRFLTIDGWKRLD

HA

ELSKLEHIALPRKLESSSLQLGLRGQYPYDVPDYAID

FIG._12B

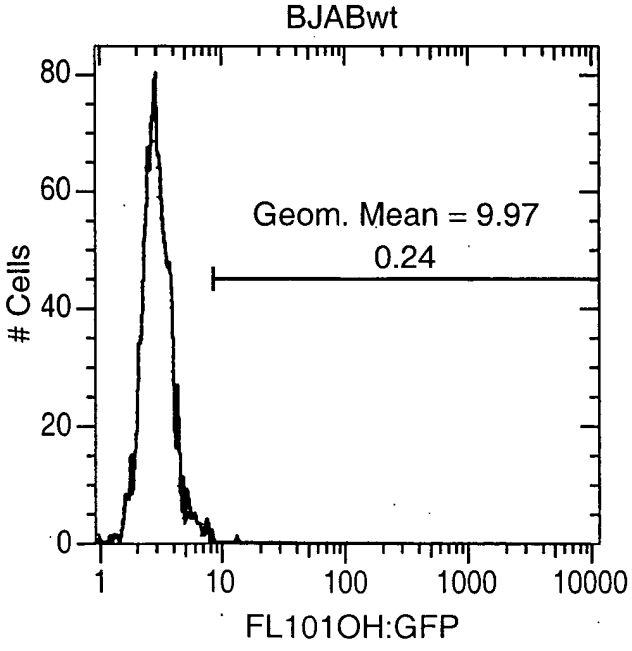


FIG._12D-1

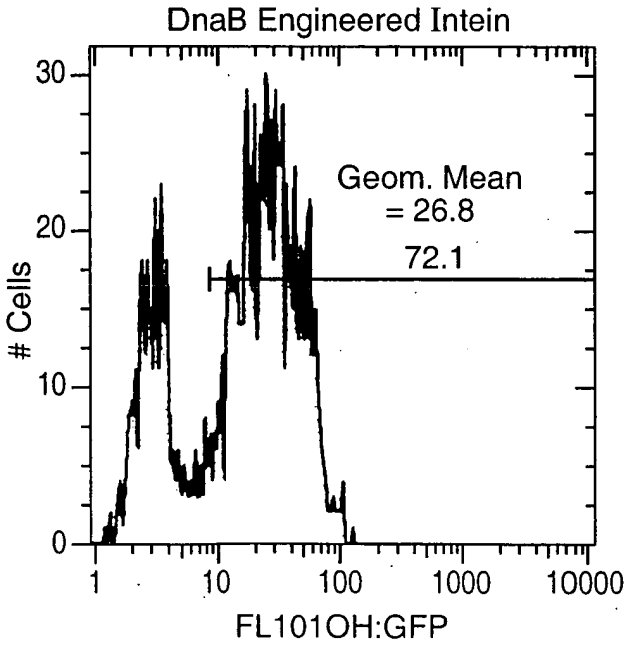


FIG._12D-2

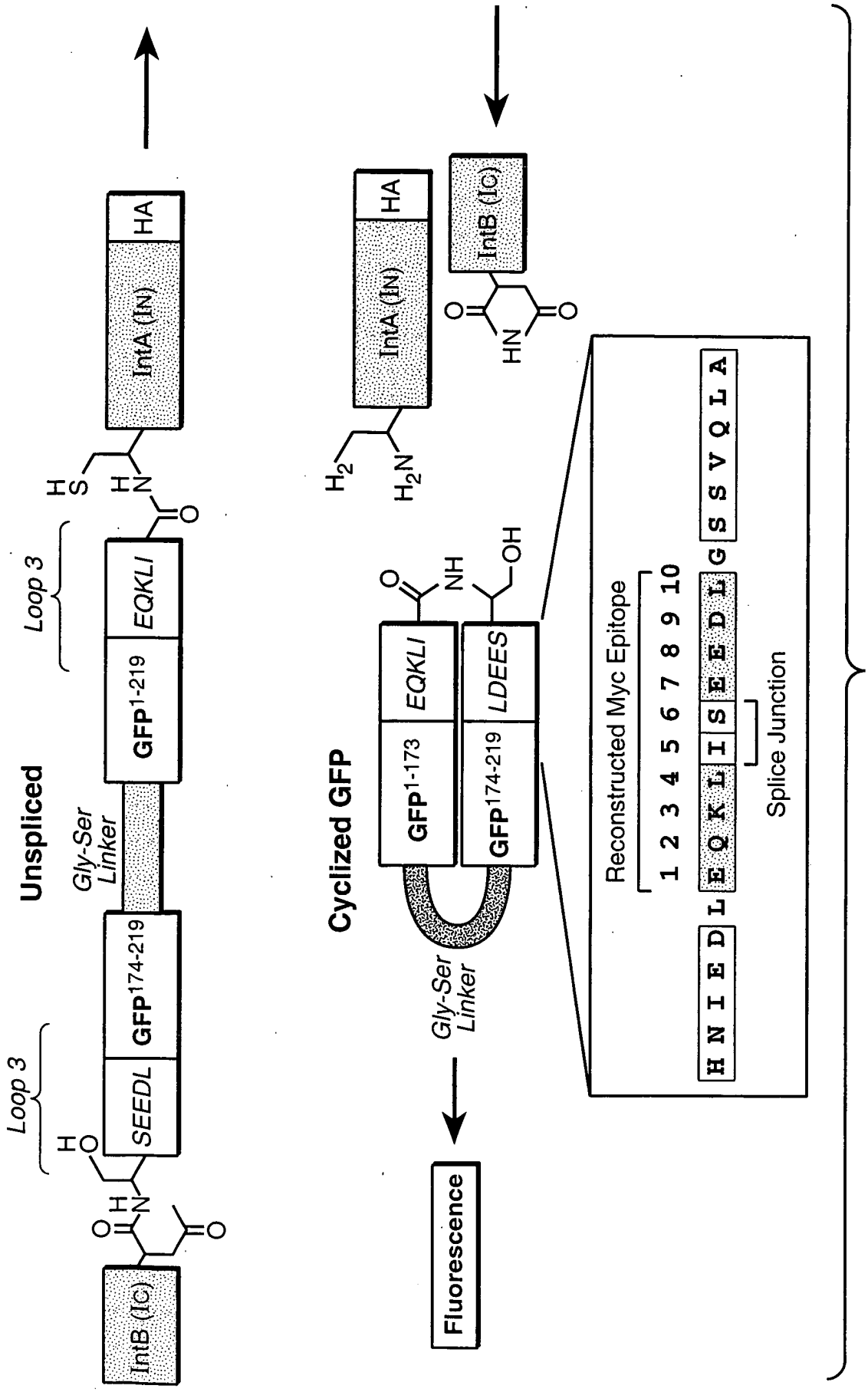

Unspliced
IntB (IC)
SEEDL
GFP174-219
Gly-Ser Linker
GFP1-219
EQKLI
IntA (IN)
HA
Loop 3
Loop 3
Cyclized GFP
GFP1-173
EQKLI
GFP174-219
LDEES
Gly-Ser Linker
Fluorescence
Reconstructed Myc Epitope
1 2 3 4 5 6 7 8 9 10
H N I E D L E Q K L I S E E D L G S S V Q L A
Splice Junction

FIG._12C-1

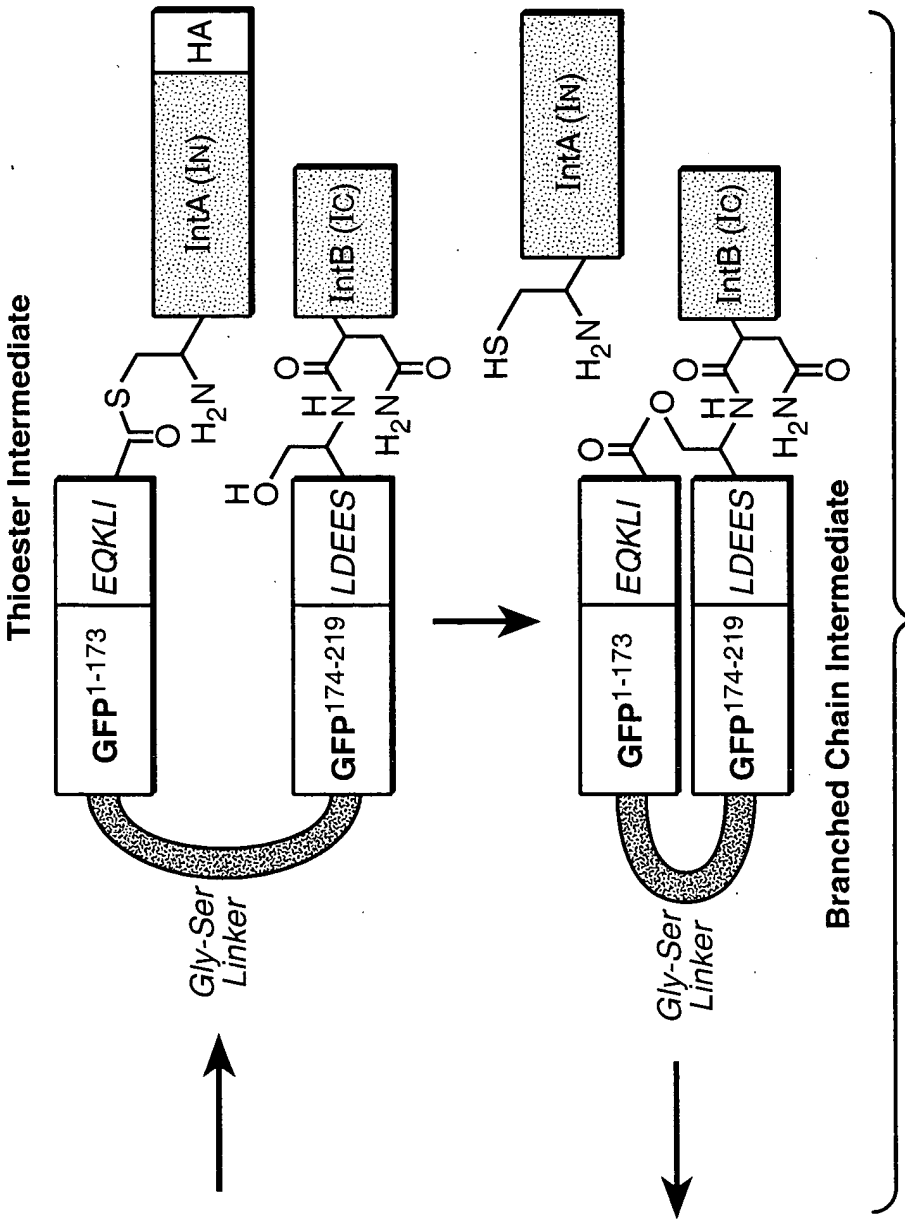

Thioester Intermediate
GFP1-173
EQKLI
IntA (IN)
HA
IntB (IC)
LDEES
GFP174-219
Gly-Ser Linker
Branched Chain Intermediate

FIG._12C-2

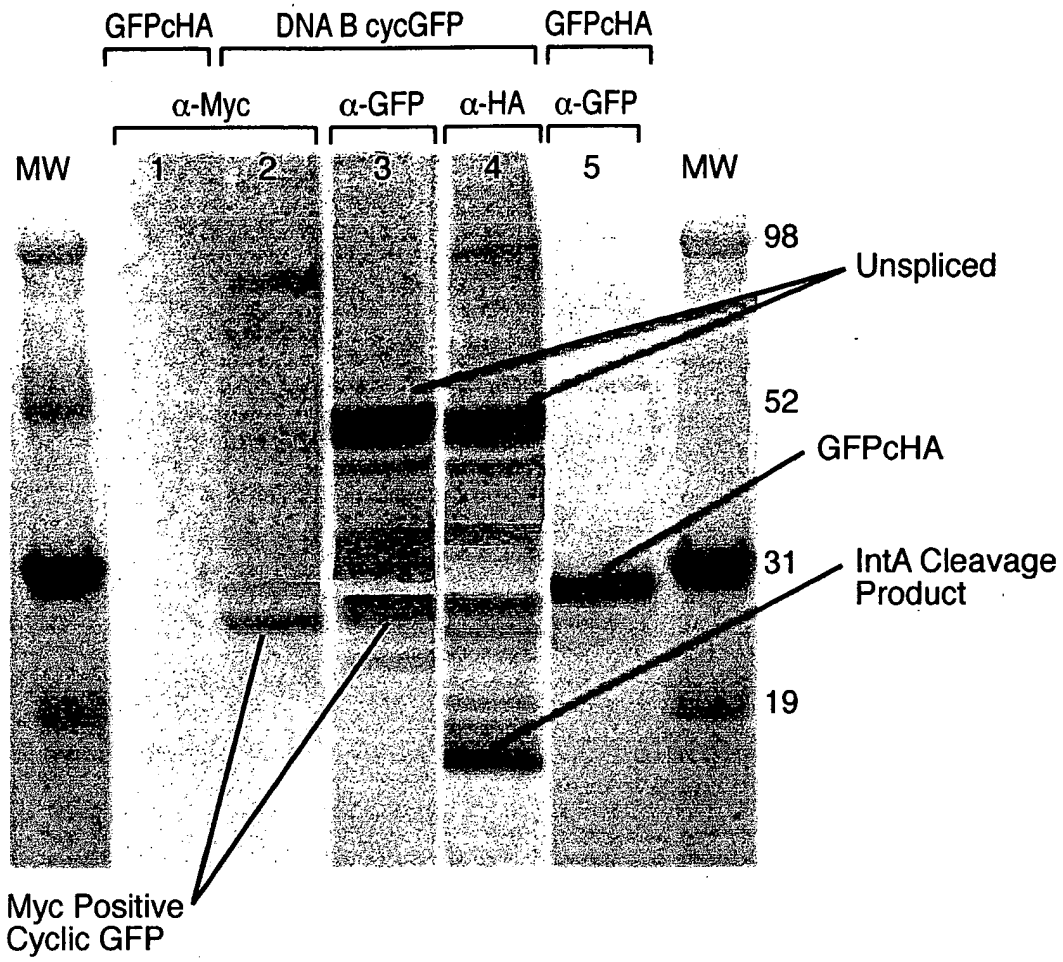


FIG._12E

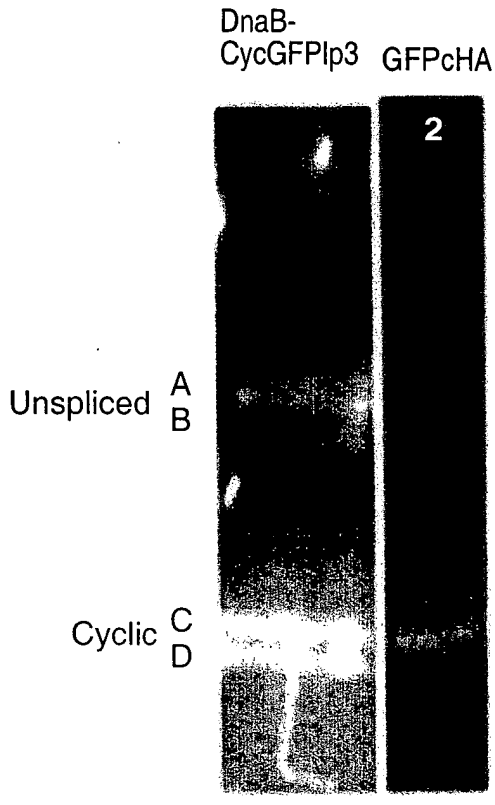


FIG._12F

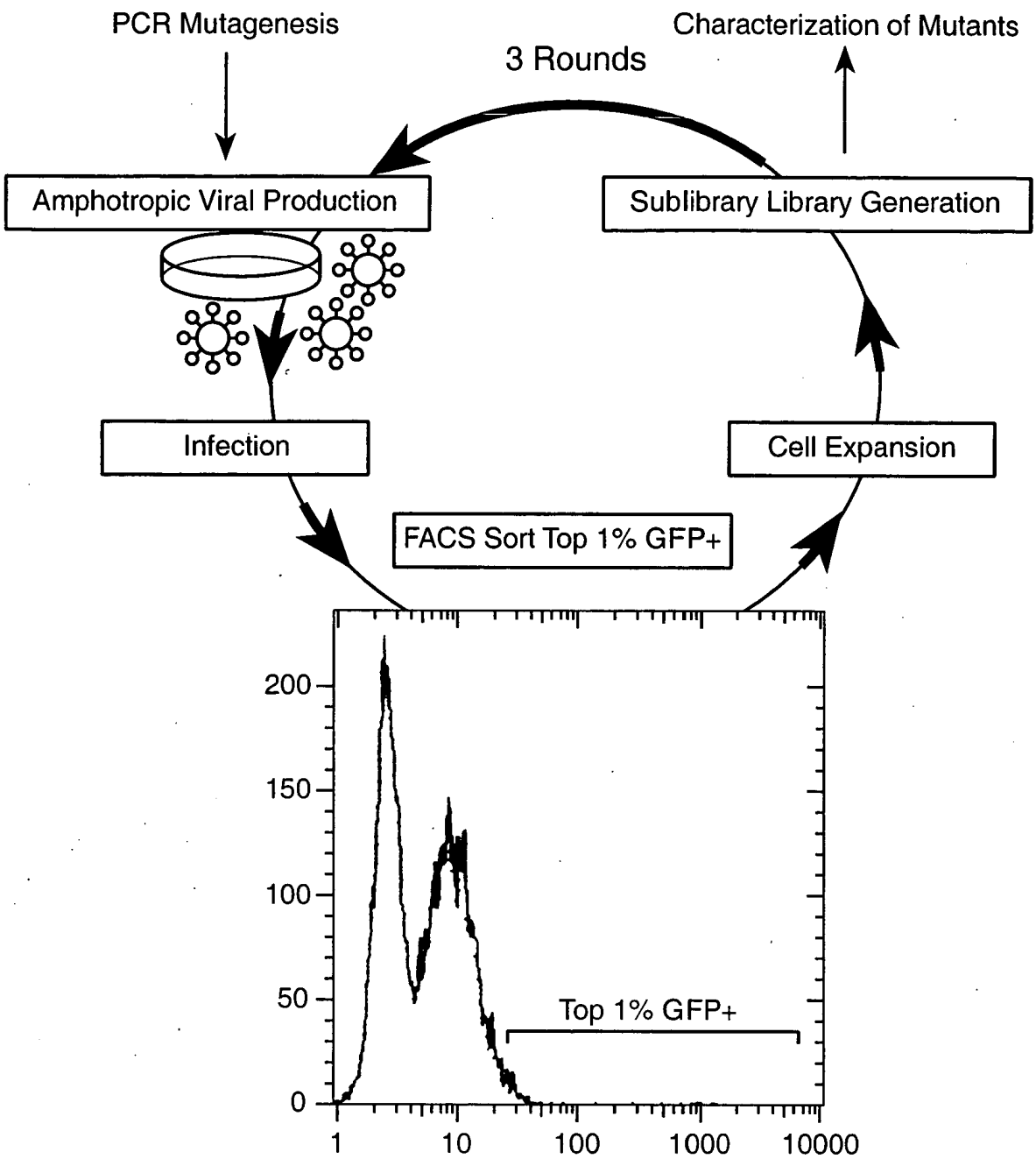

PCR Mutagenesis
3 Rounds
Characterization of Mutants
Amphotropic Viral Production
Sublibrary Library Generation
Infection
Cell Expansion
FACS Sort Top 1% GFP+
200
150
100
50
0
Top 1% GFP+
1
10
100
1000
10000

FIG._13A

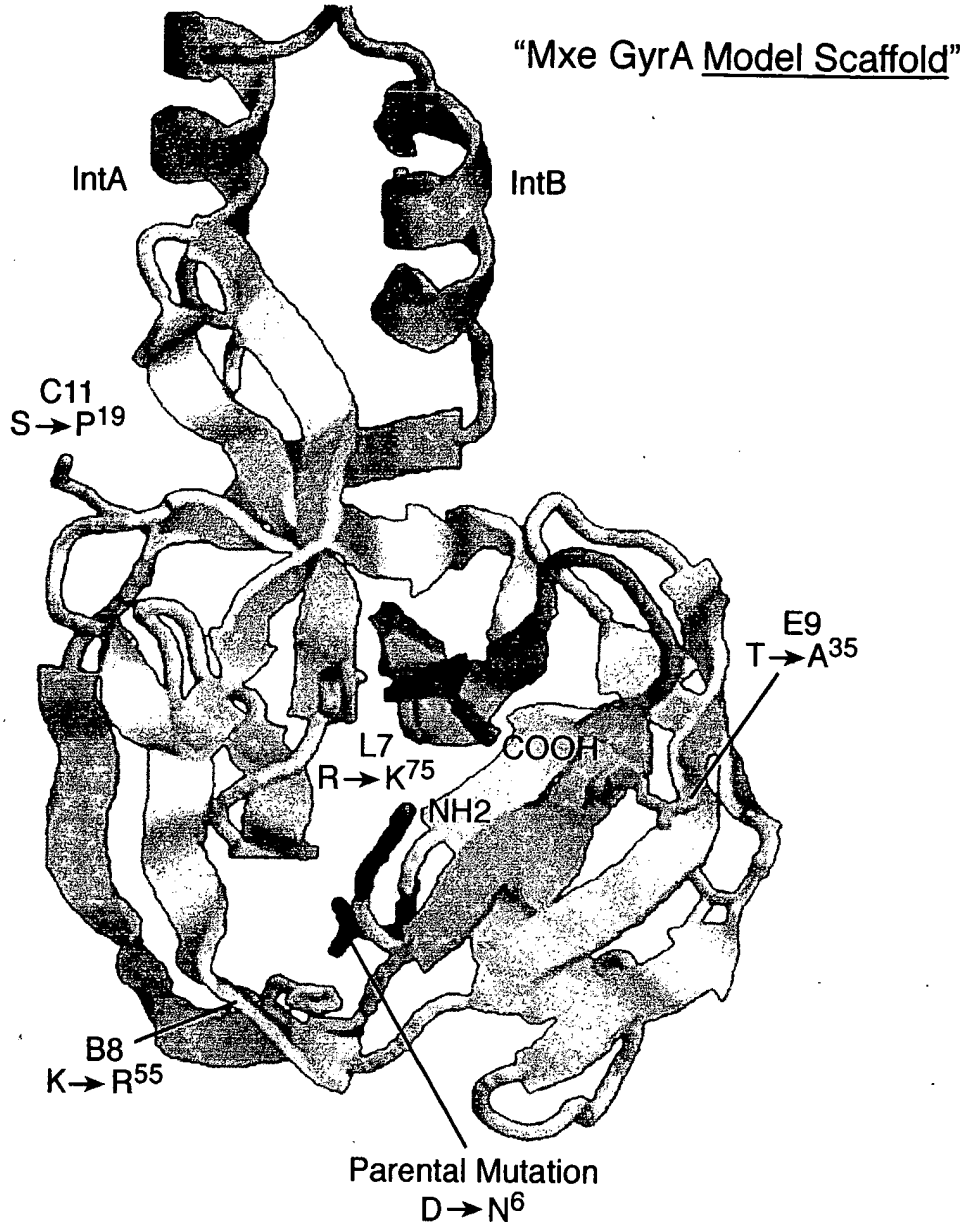

"Mxe GyrA Model Scaffold"
IntA
IntB
C11
S→P19
E9
T→A35
L7
R→K75
COOH
NH2
B8
K→R55
Parental Mutation
D→N6

FIG._13B

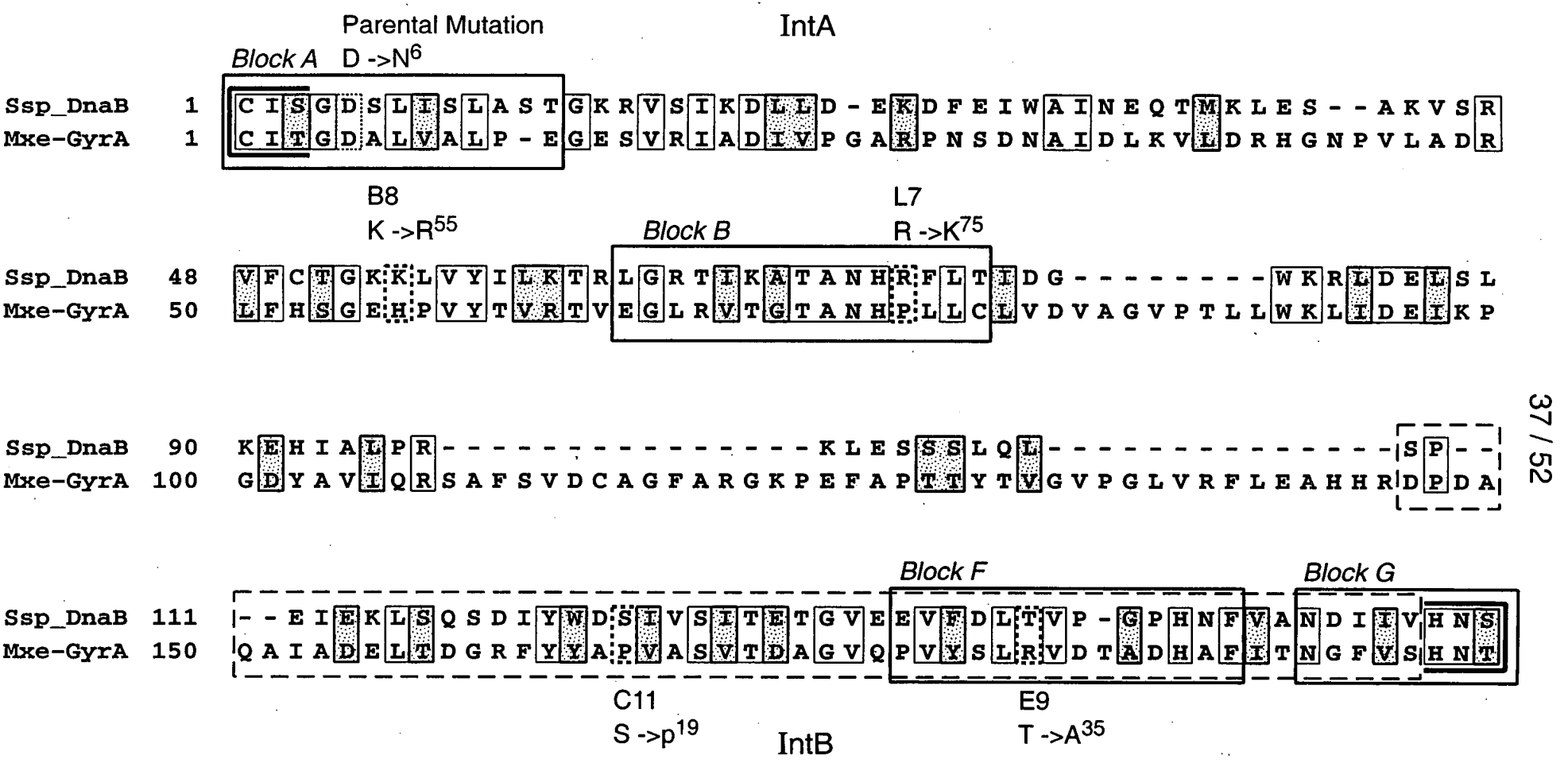
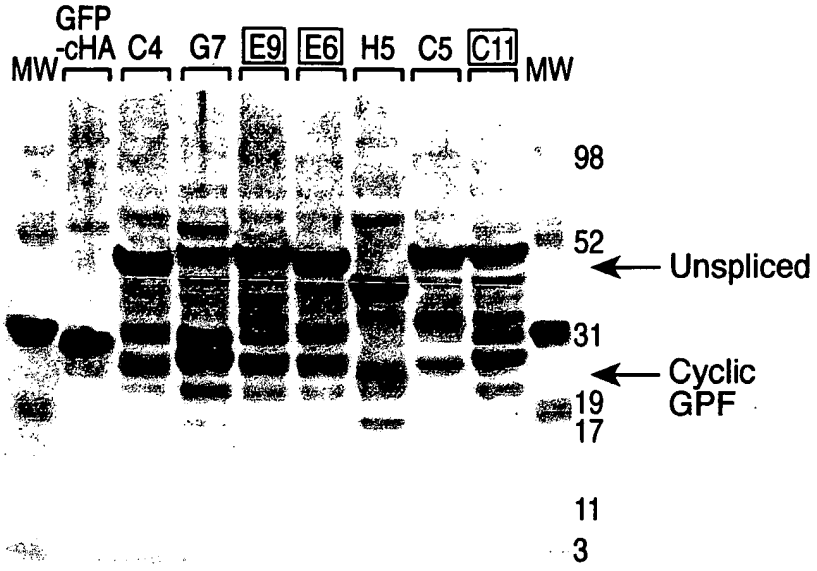
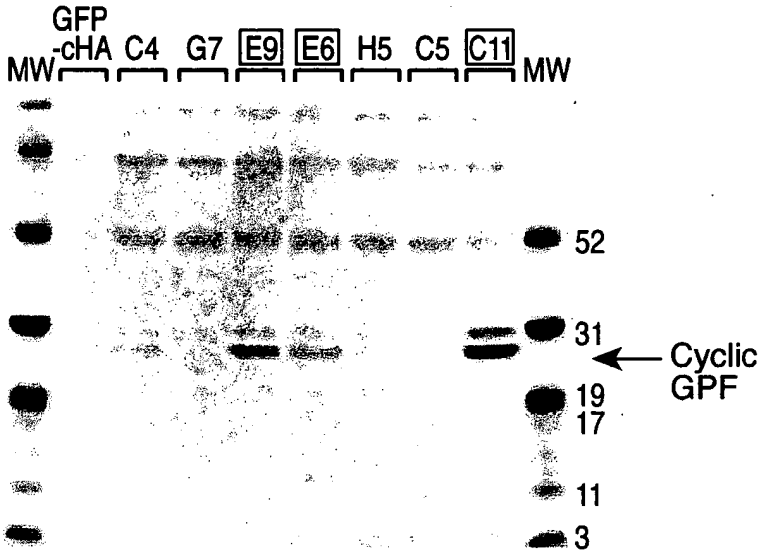

Parental Mutation
IntA
Block A
D ->N6
Ssp_DnaB 1 C I S G D S L I S L A S T G K R V S I K D L L D - E K D F E I W A I N E Q T M K L E S - - A K V S R
Mxe-GyrA 1 C I T G D A L V A L P - E G E S V R I A D I V P G A R P N S D N A I D L K V L D R H G N P V L A D R
B8
K ->R55
Block B
L7
R ->K75
Ssp_DnaB 48 V F C T G K K L V Y I L K T R L G R T I K A T A N H R F L T I D G - - - - - - - - W K R L D E L S L
Mxe-GyrA 50 L F H S G E H P V Y T V R T V E G L R V T G T A N H P L L C L V D V A G V P T L L W K L I D E I K P
Ssp_DnaB 90 K E H I A L P R - - - - - - - - - - - - - - - K L E S S S L Q L - - - - - - - - - - - - - - S P - -
Mxe-GyrA 100 G D Y A V I Q R S A F S V D C A G F A R G K P E F A P T T Y T V G V P G L V R F L E A H H R D P D A
Block F
Block G
Ssp_DnaB 111 - - E I E K L S Q S D I Y W D S I V S I T E T G V E E V F D L T V P - G P H N F V A N D I I V H N S
Mxe-GyrA 150 Q A I A D E L T D G R F Y Y A P V A S V T D A G V Q P V Y S L R V D T A D H A F I T N G F V S H N T
C11
S ->p19
IntB
E9
T ->A35

FIG._13C



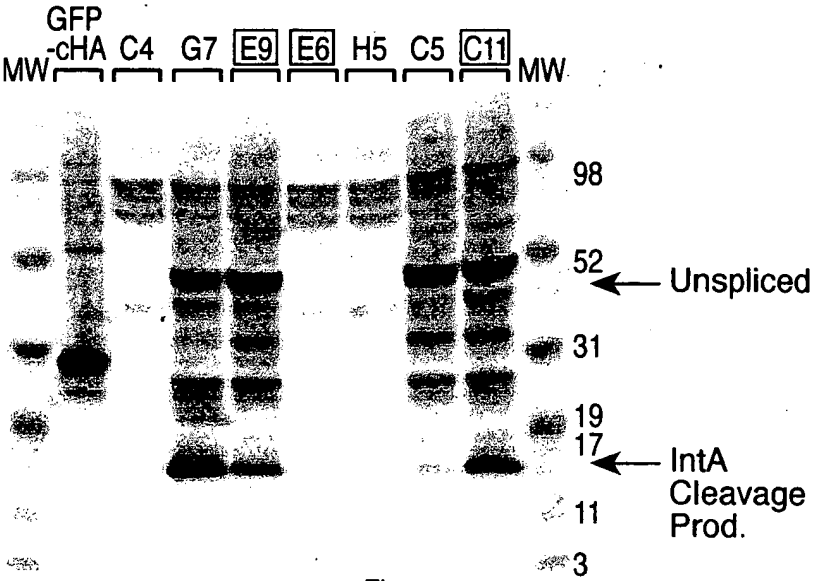
α-GFP
Identifies All Splice Intermediates and Final Product

FIG._13D-1



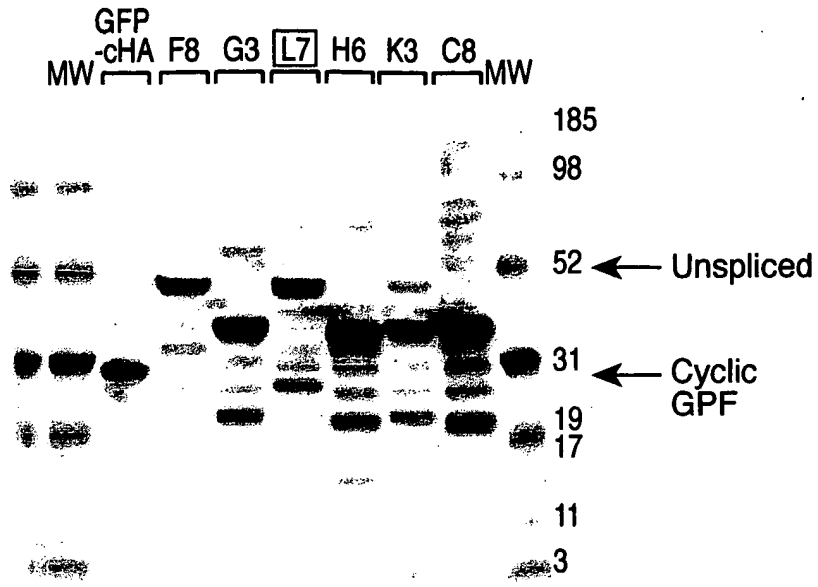
α-myc
Identifies Only Cyclic Product with Reconstructed Myc Epitope

FIG._13D-2



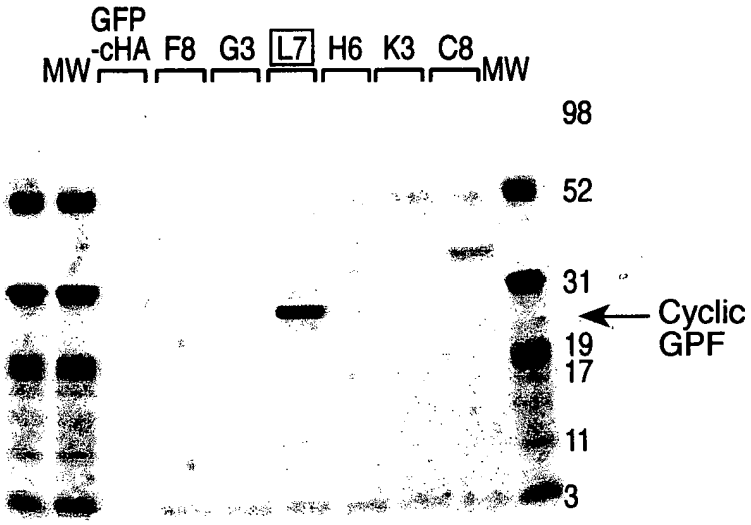
α-Flu
Identifies IntA Splice Intermediates and Cleavage Product

FIG._13D-3



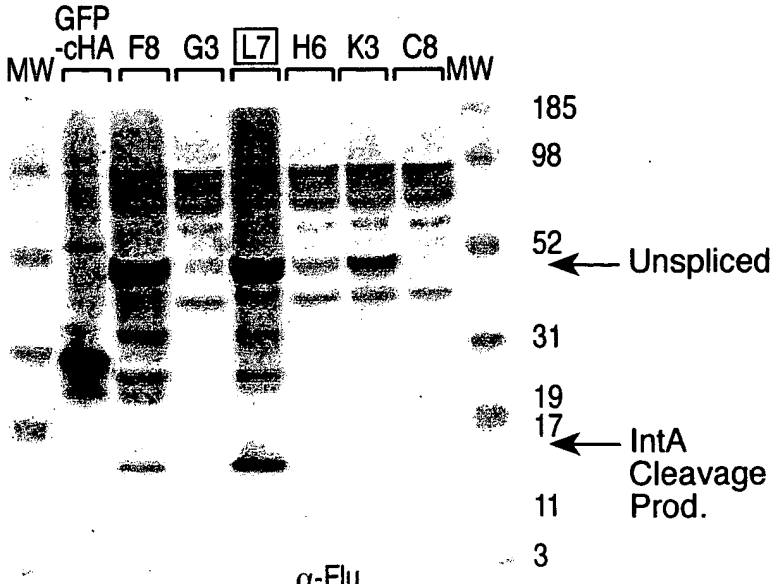
α-GFP
Identifies All Splice Intermediates and Final Product

FIG._13D-4



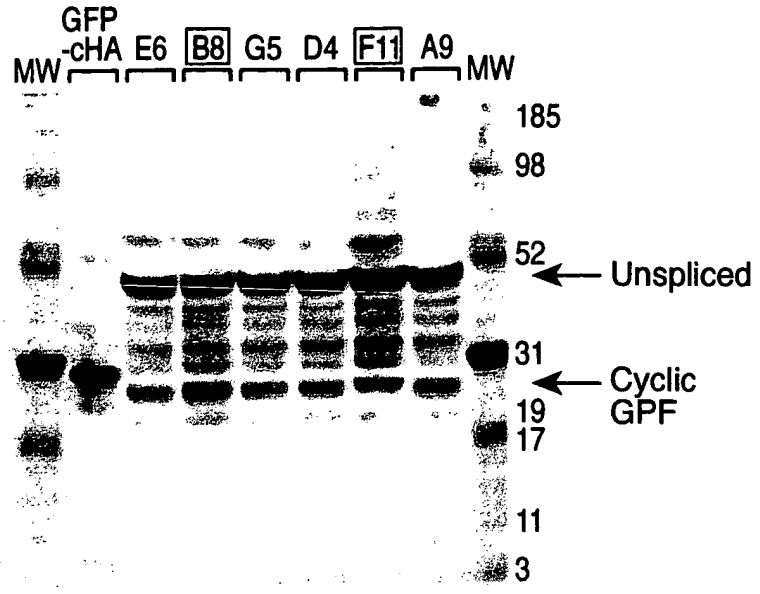
α-myc
Identifies Only Cyclic Product with Reconstructed Myc Epitope

FIG._13D-5



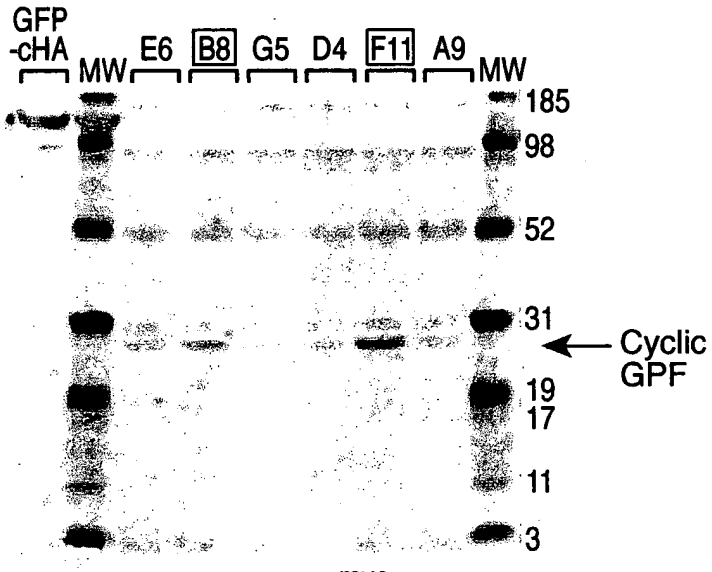
α-Flu
Identifies IntA Splice Intermediates and Cleavage Product

FIG._13D-6



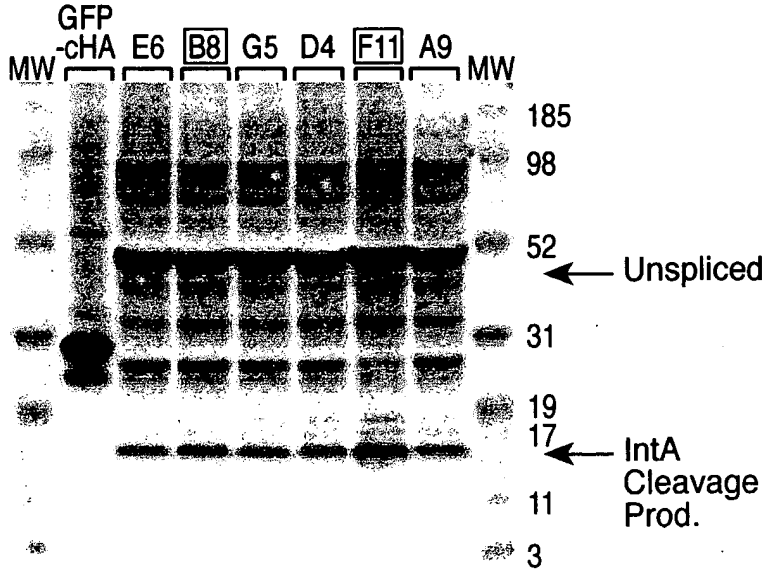
α-GFP
Identifies All Splice Intermediates and Final Product

FIG._13D-7



α-myc
Identifies Only Cyclic Product with Reconstructed Myc Epitope

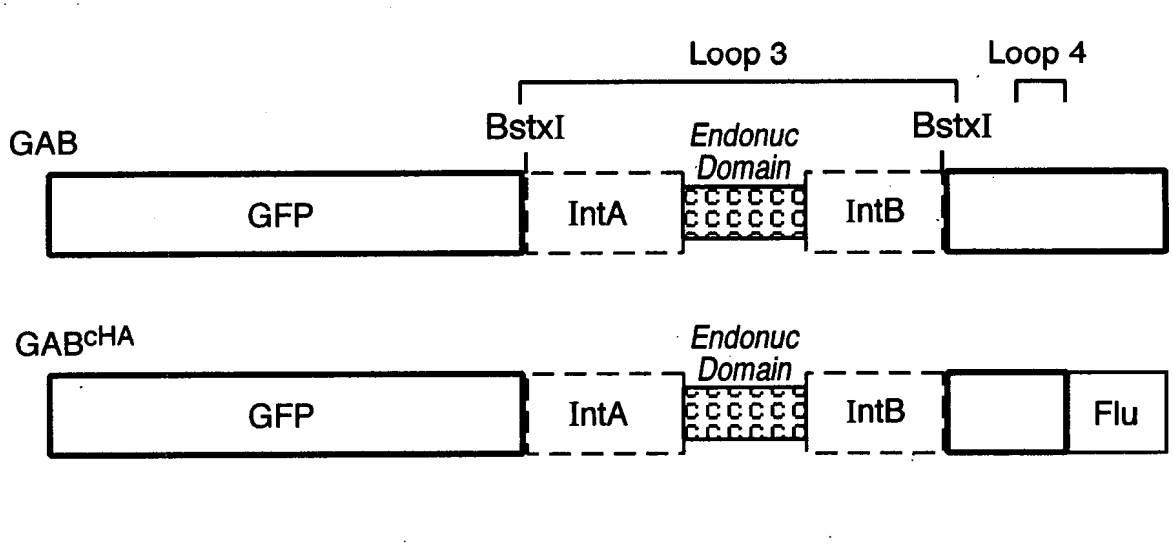
FIG._13D-8



α-Flu
Identifies IntA Splice Intermediates and Cleavage Product

FIG._13D-9

FIG._14A



Transfected PhxA Cells

GAB

GAB-cHA

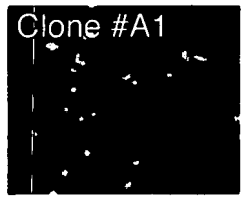
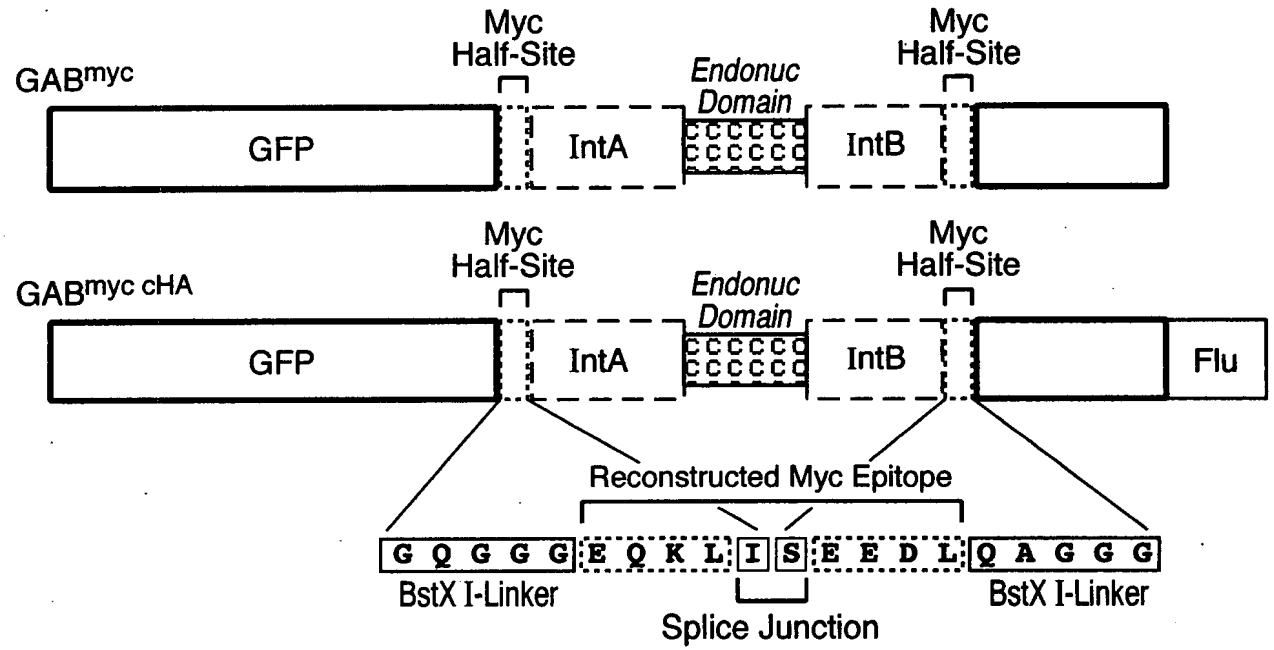
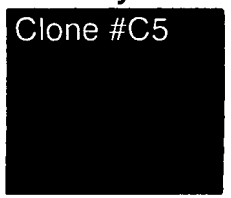


FIG._14B



GAB-Myc



GAB-Mtc-cHA

FIG._14C-1

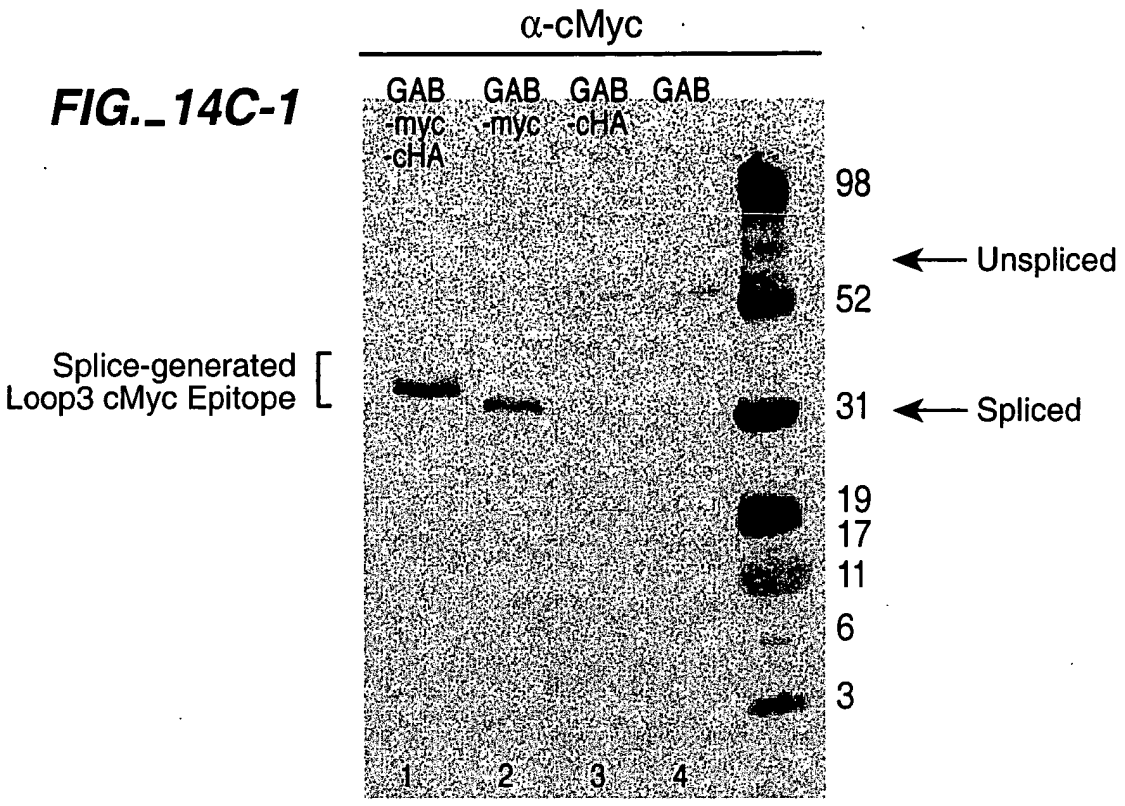


FIG._14C-2

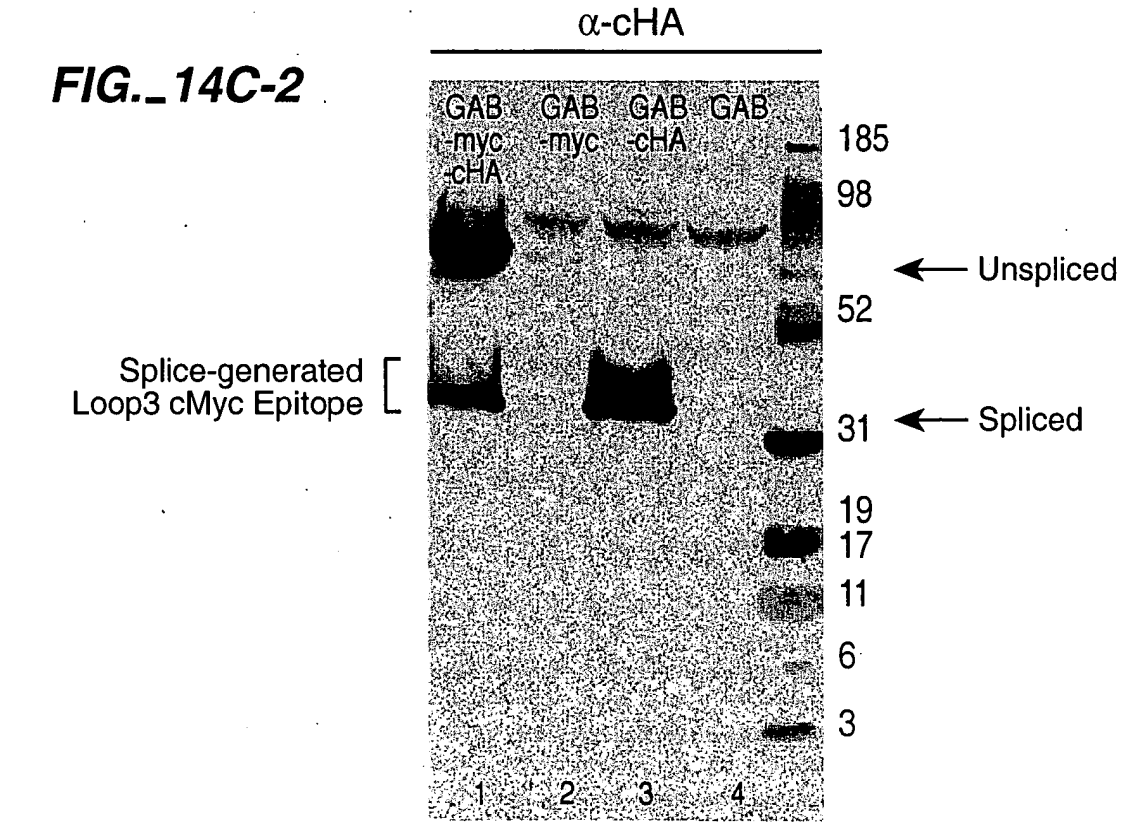
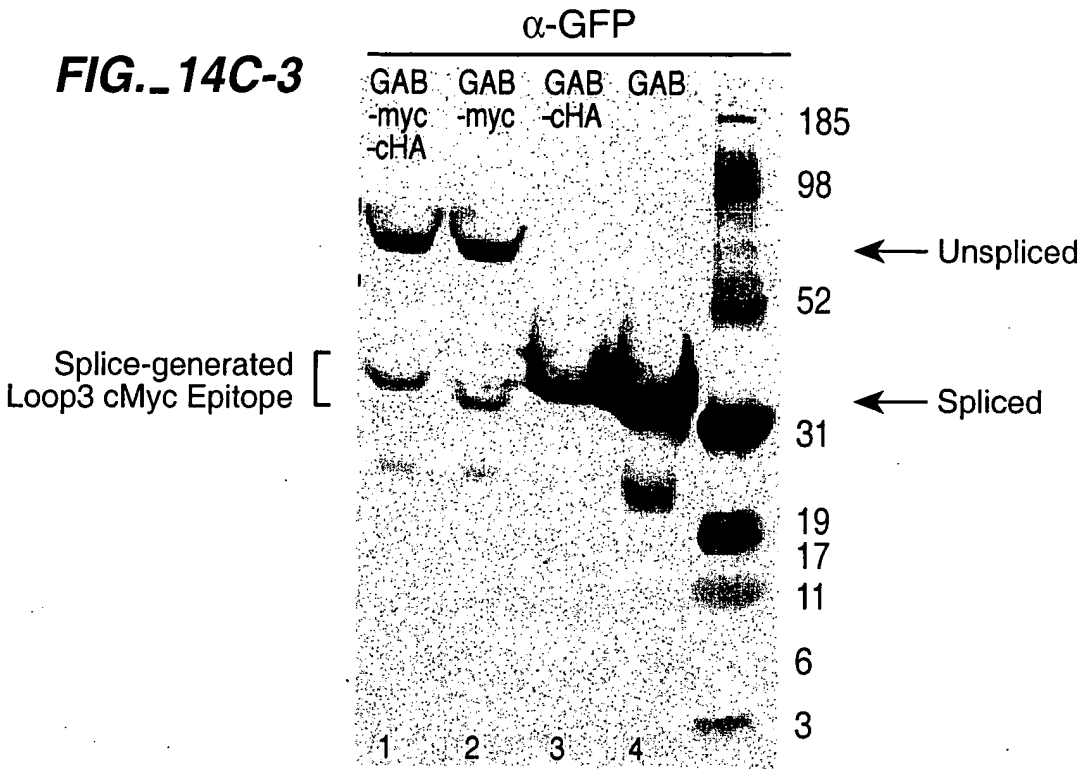


FIG._14C-3


α-GFP
GAB -myc -cHA
GAB -myc
GAB -cHA
GAB
185
98
52
31
19
17
11
6
3
1
2
3
4
Unspliced
Spliced
Splice-generated Loop3 cMyc Epitope

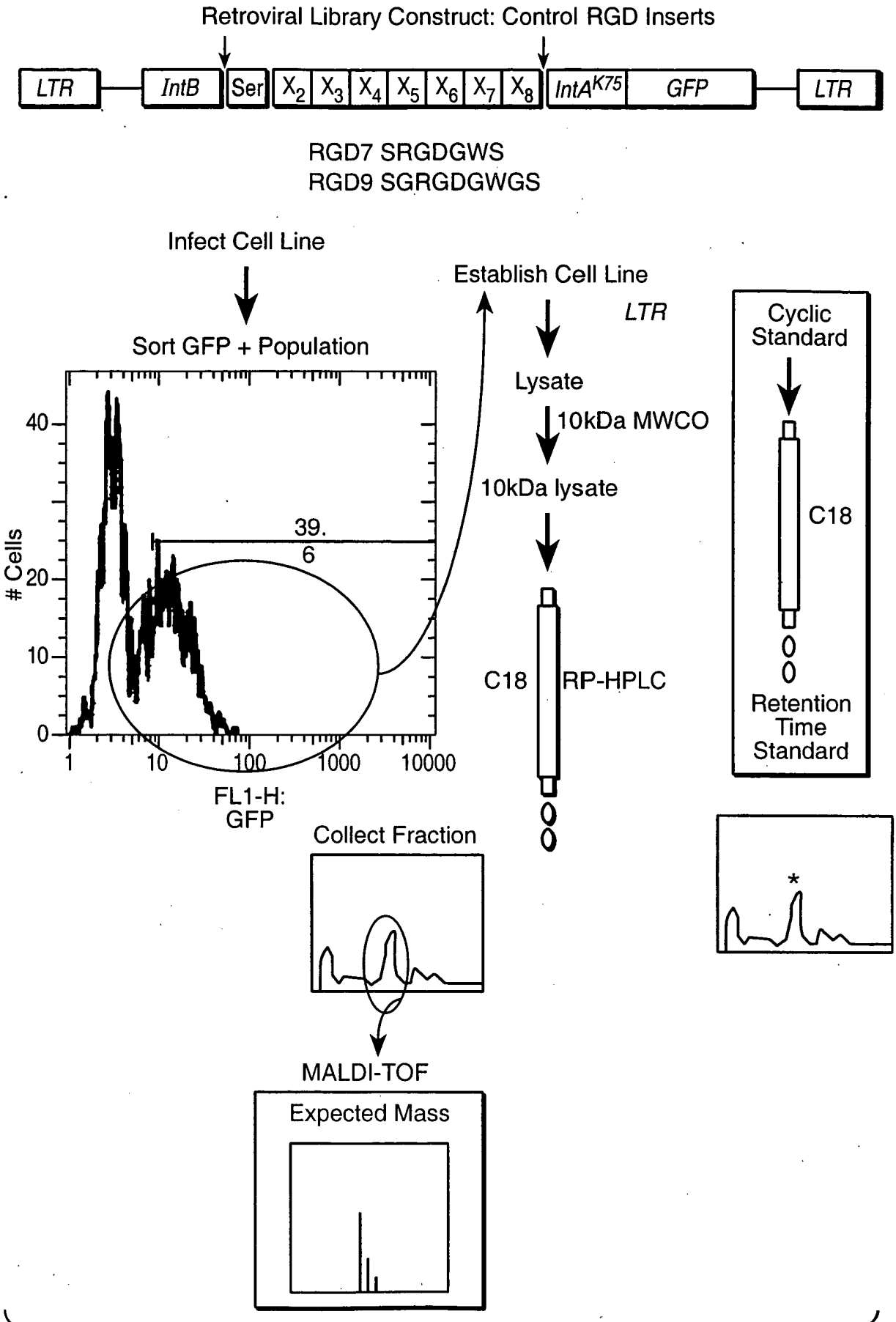

Retroviral Library Construct: Control RGD Inserts
LTR
IntB
Ser
X2 X3 X4 X5 X6 X7 X8
IntA K75
GFP
LTR
RGD7 SRGDGWS
RGD9 SGRGDGWGS
Infect Cell Line
Sort GFP + Population
Establish Cell Line
LTR
Lysate
10kDa MWCO
10kDa lysate
C18
RP-HPLC
Cyclic Standard
C18
Retention Time Standard
Cells
39.6
FL1-H: GFP
Collect Fraction
MALDI-TOF
Expected Mass

FIG._15A

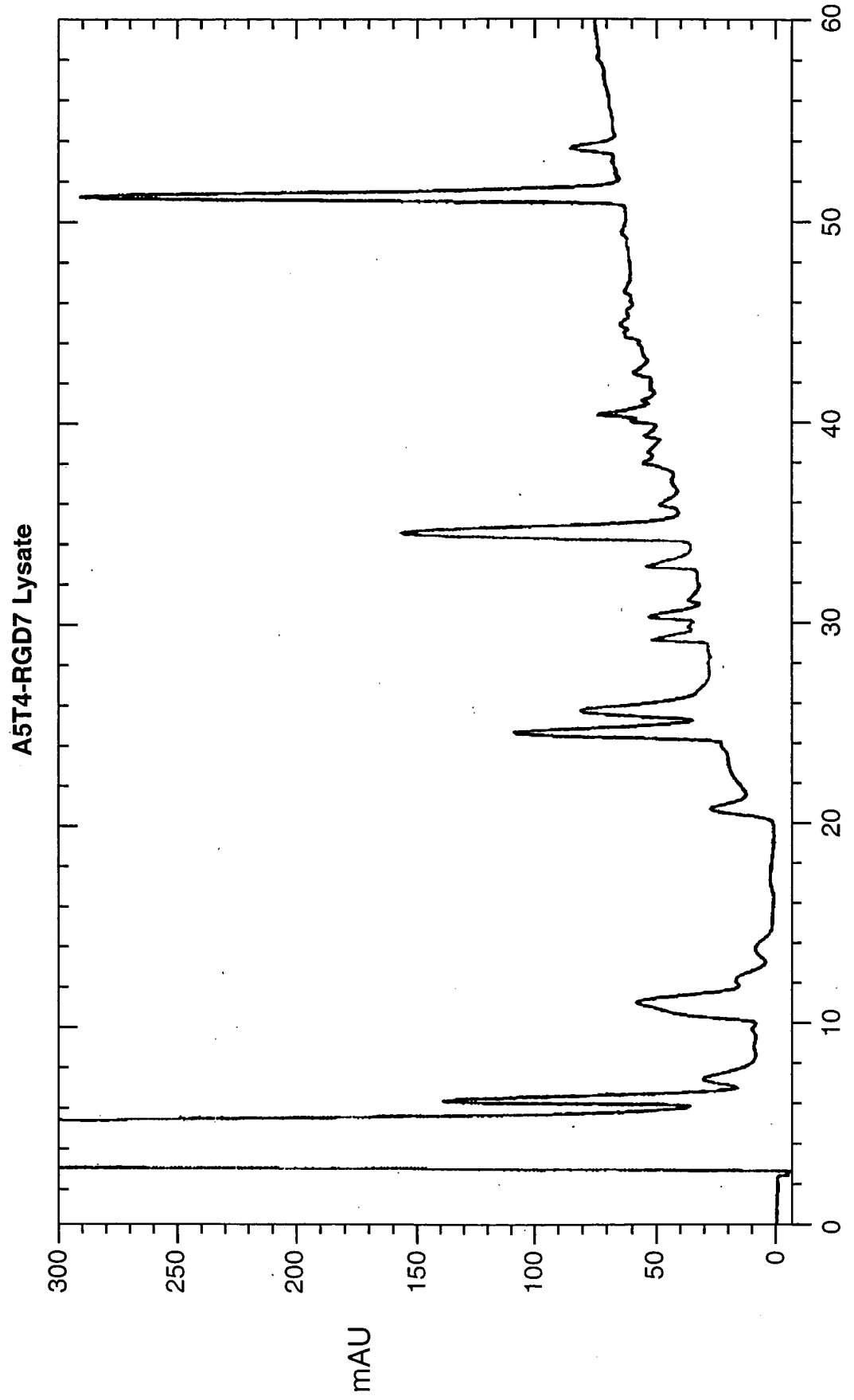

A5T4-RGD7 Lysate
mAU
300
250
200
150
100
50
0
0
10
20
30
40
50
60
Min.

FIG._15B-1

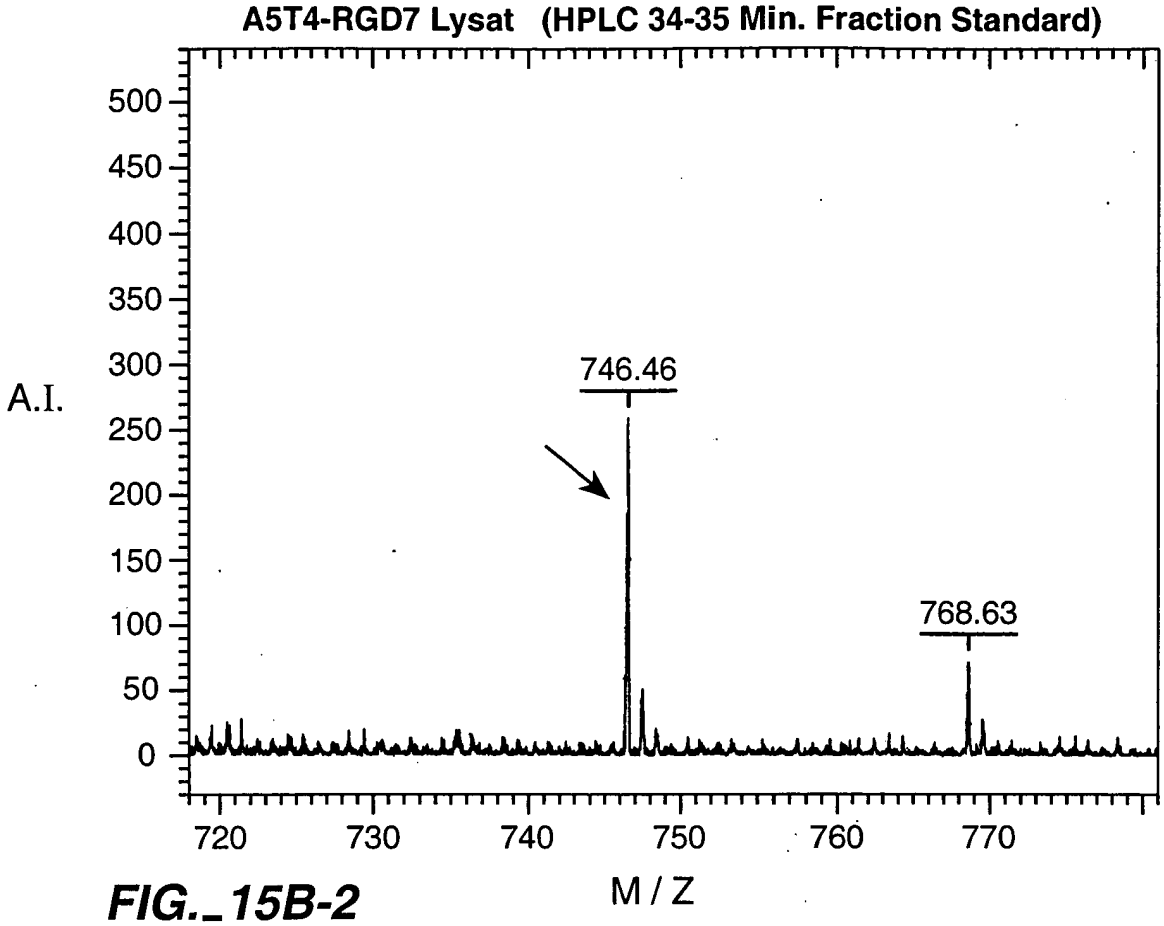


FIG._15B-2

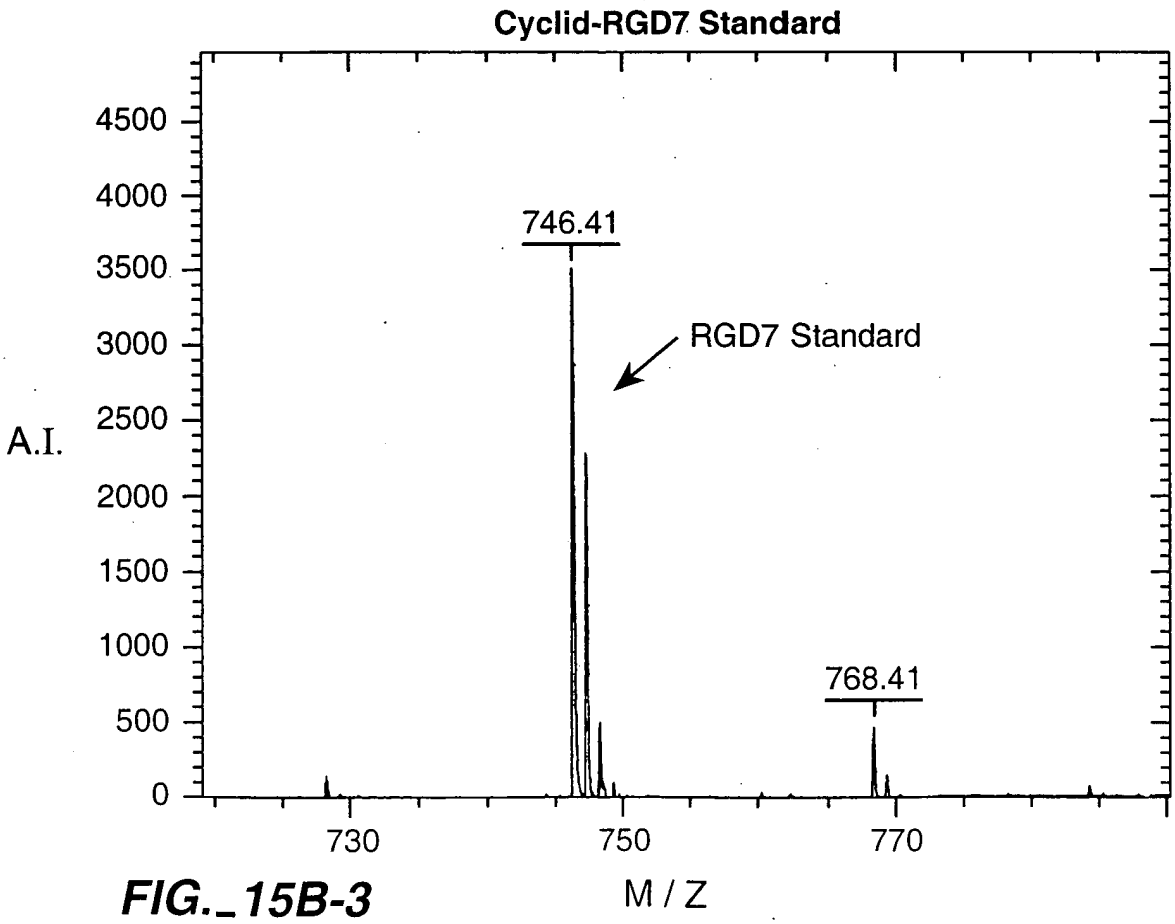


FIG._15B-3

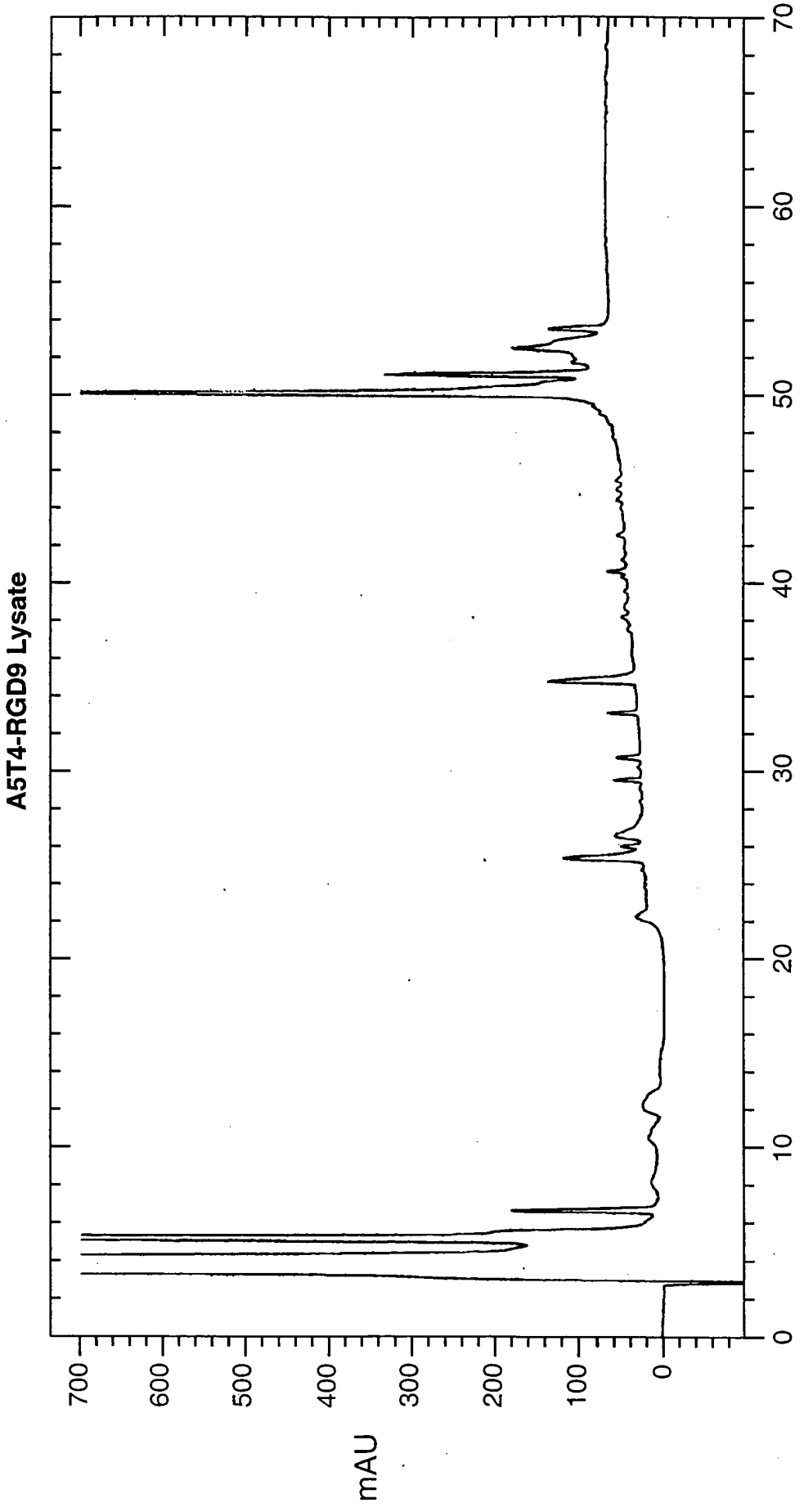

A5T4-RGD9 Lysate
mAU
700
600
500
400
300
200
100
0
0
10
20
30
40
50
60
70
Min.

FIG._15C-1

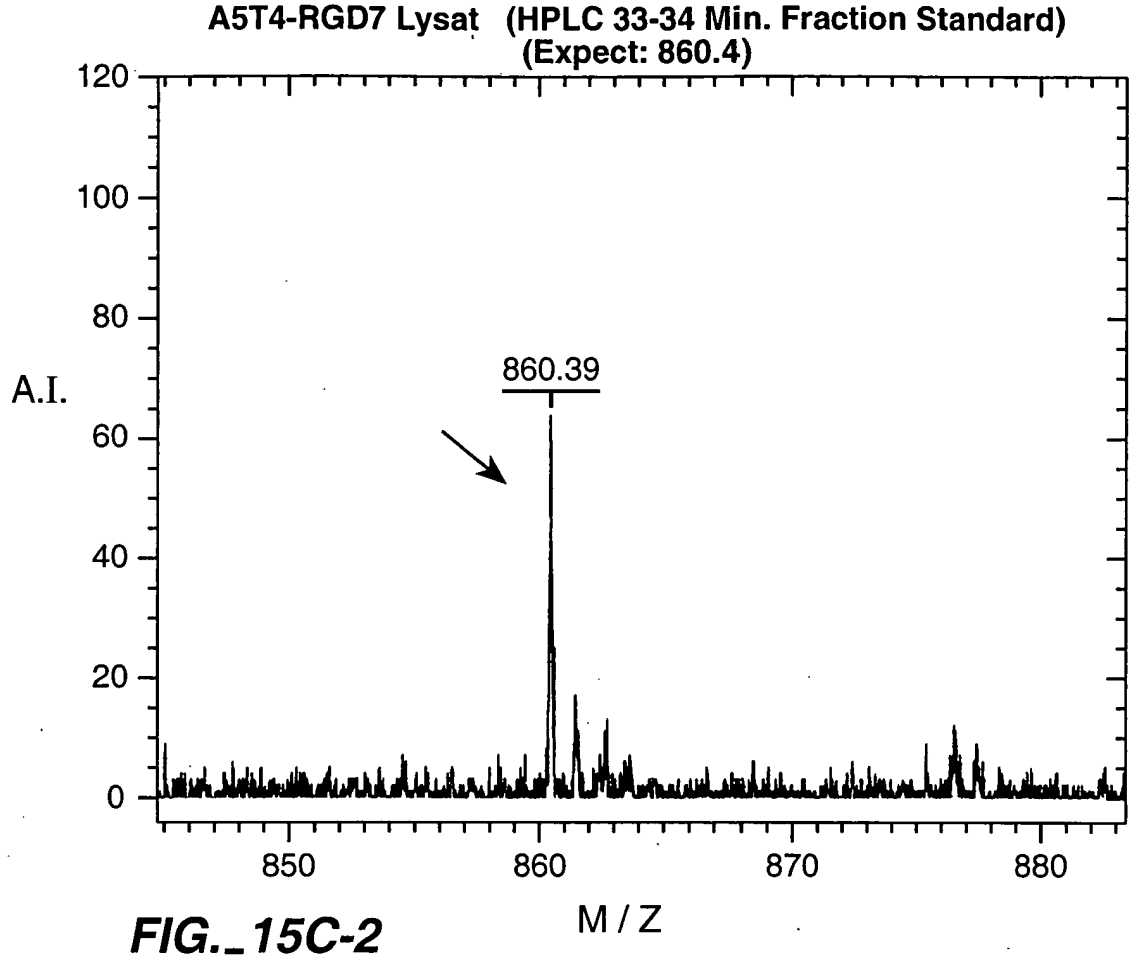


FIG._15C-2

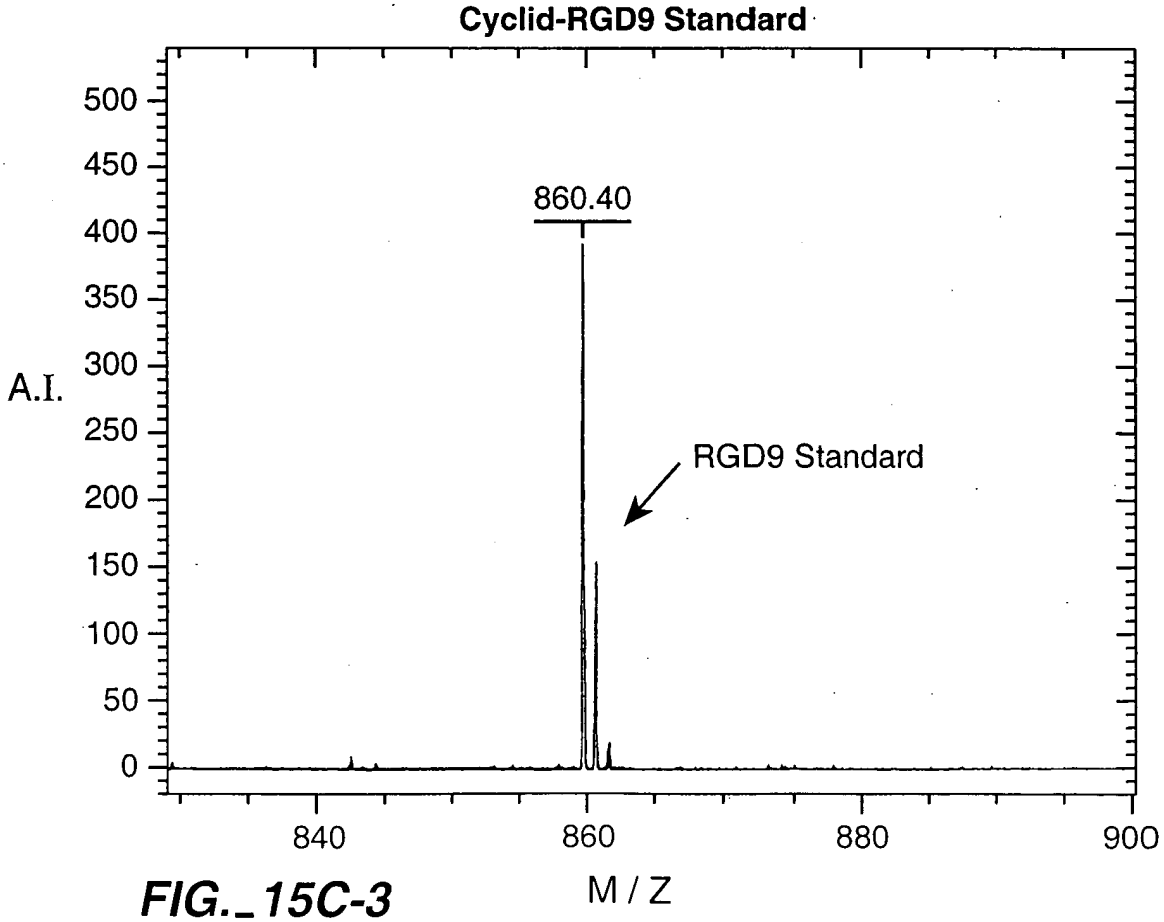


FIG._15C-3

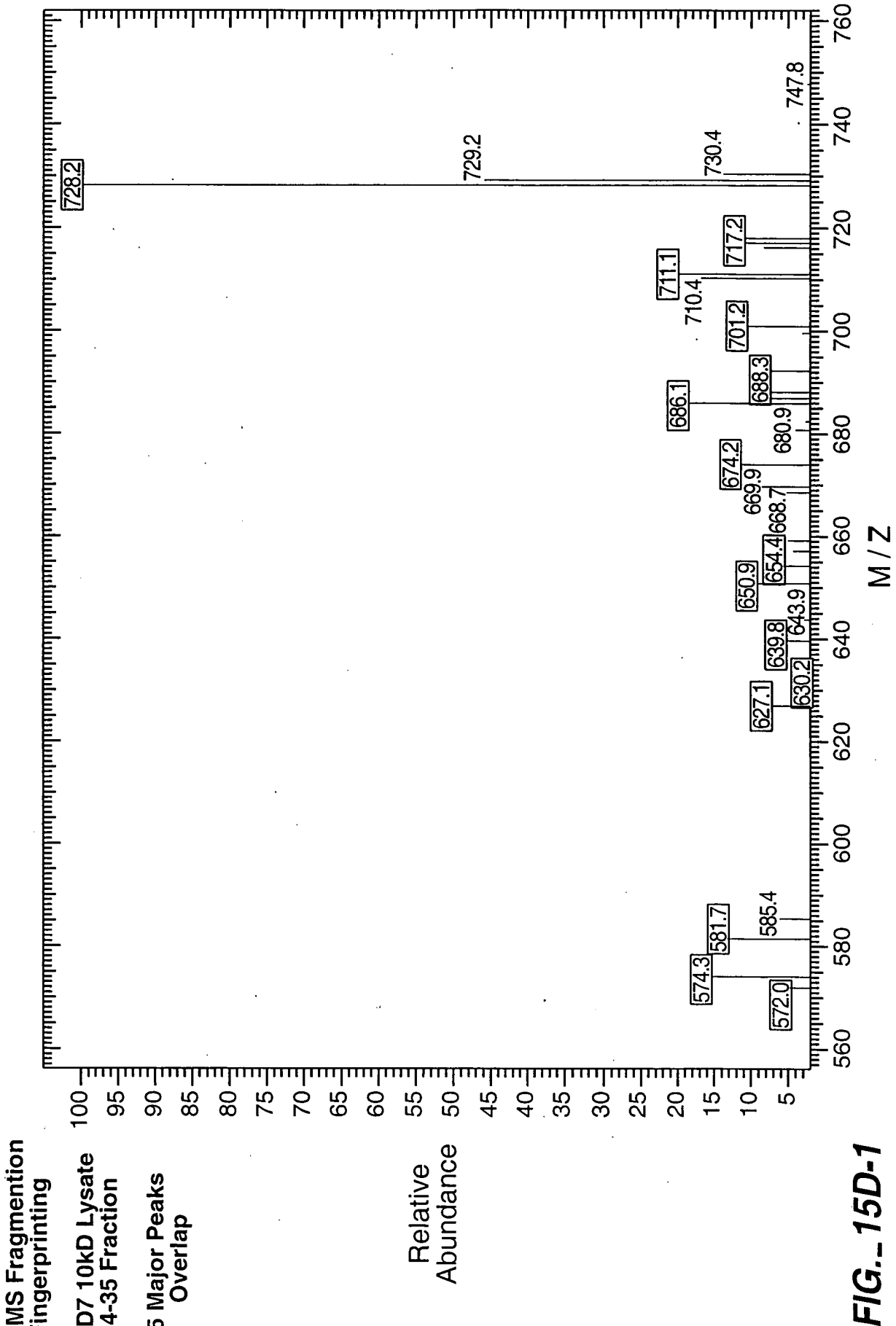

LC / MS Fragmention Fingerprinting
RGD7 10kD Lysate 34-35 Fraction
15 Major Peaks Overlap
Relative Abundance
M / Z
728.2
729.2
730.4
747.8
711.1
710.4
717.2
701.2
688.3
686.1
680.9
674.2
669.9
668.7
654.4
650.9
643.9
639.8
630.2
627.1
585.4
581.7
574.3
572.0

FIG._15D-1

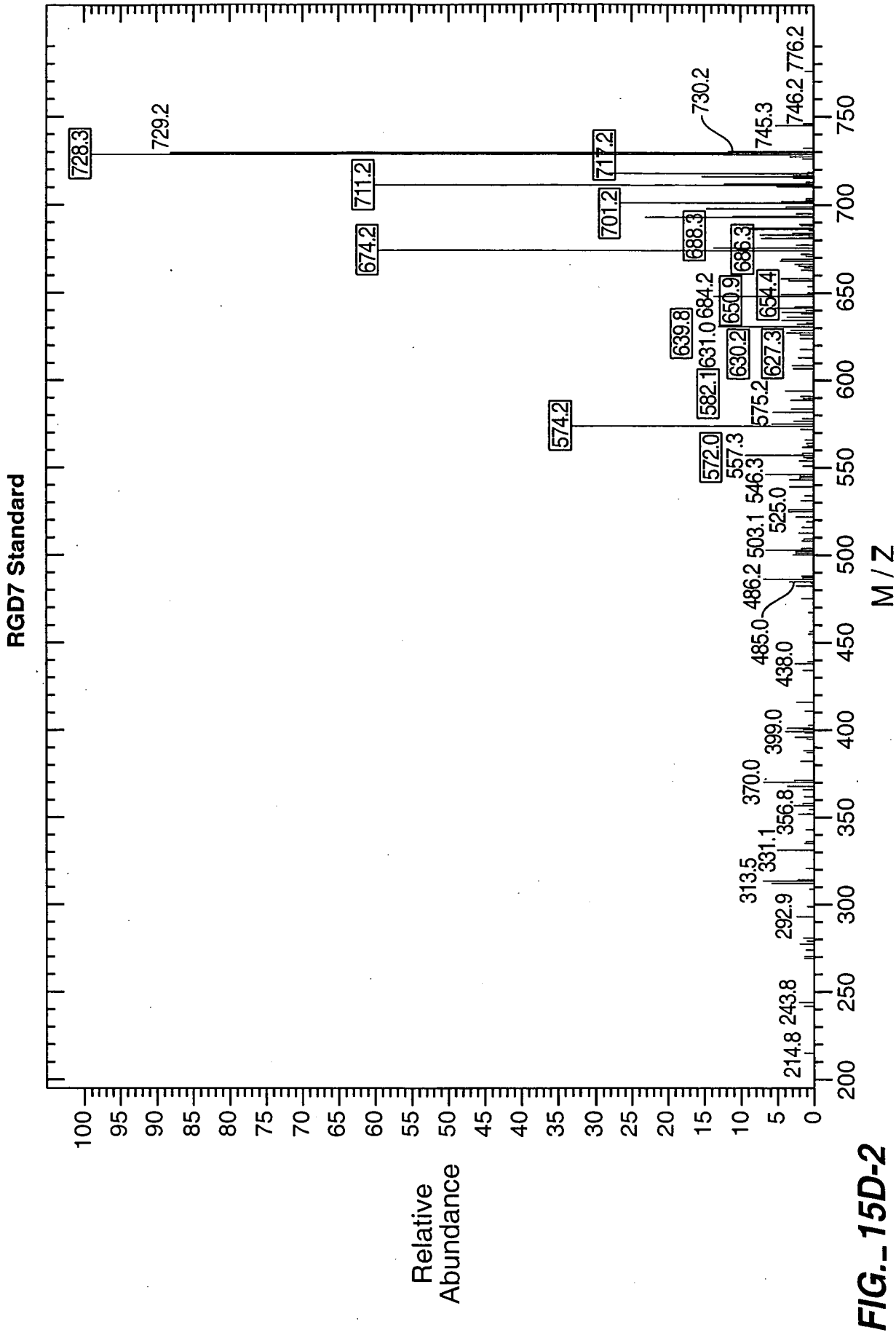

RGD7 Standard
Relative Abundance
M / Z
728.3
729.2
711.2
674.2
730.2
717.2
701.2
688.3
686.3
684.2
650.9
654.4
639.8
631.0
630.2
627.3
582.1
575.2
574.2
572.0
557.3
546.3
525.0
503.1
486.2
485.0
438.0
399.0
370.0
356.8
331.1
313.5
292.9
243.8
214.8
745.3
746.2
776.2

FIG._15D-2

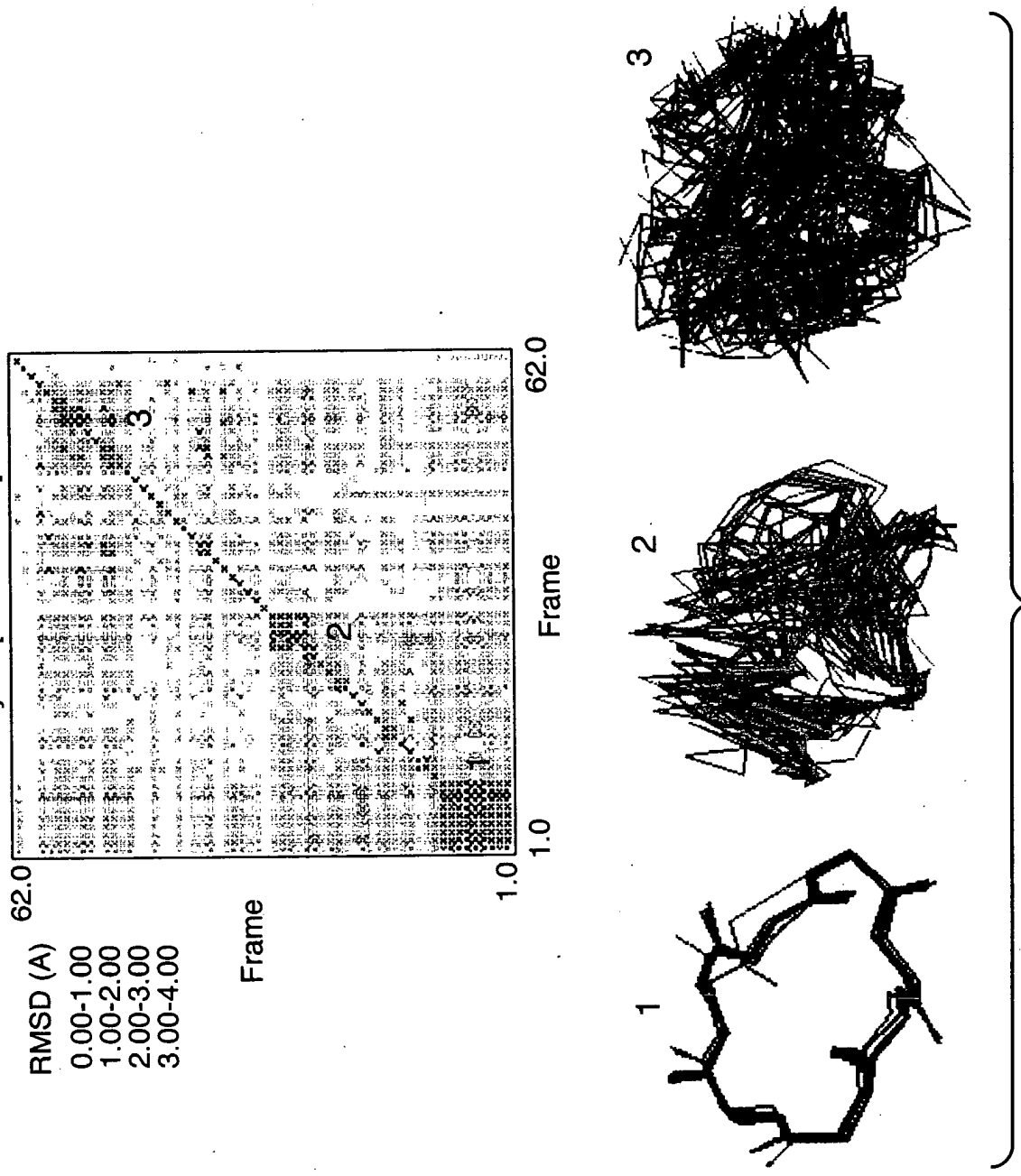

Cyclic [SRGDGWS]
RMSD (A)
0.00-1.00
1.00-2.00
2.00-3.00
3.00-4.00
62.0
1.0
Frame
1.0
62.0
Frame
1
2
3
1
2
3

FIG._16

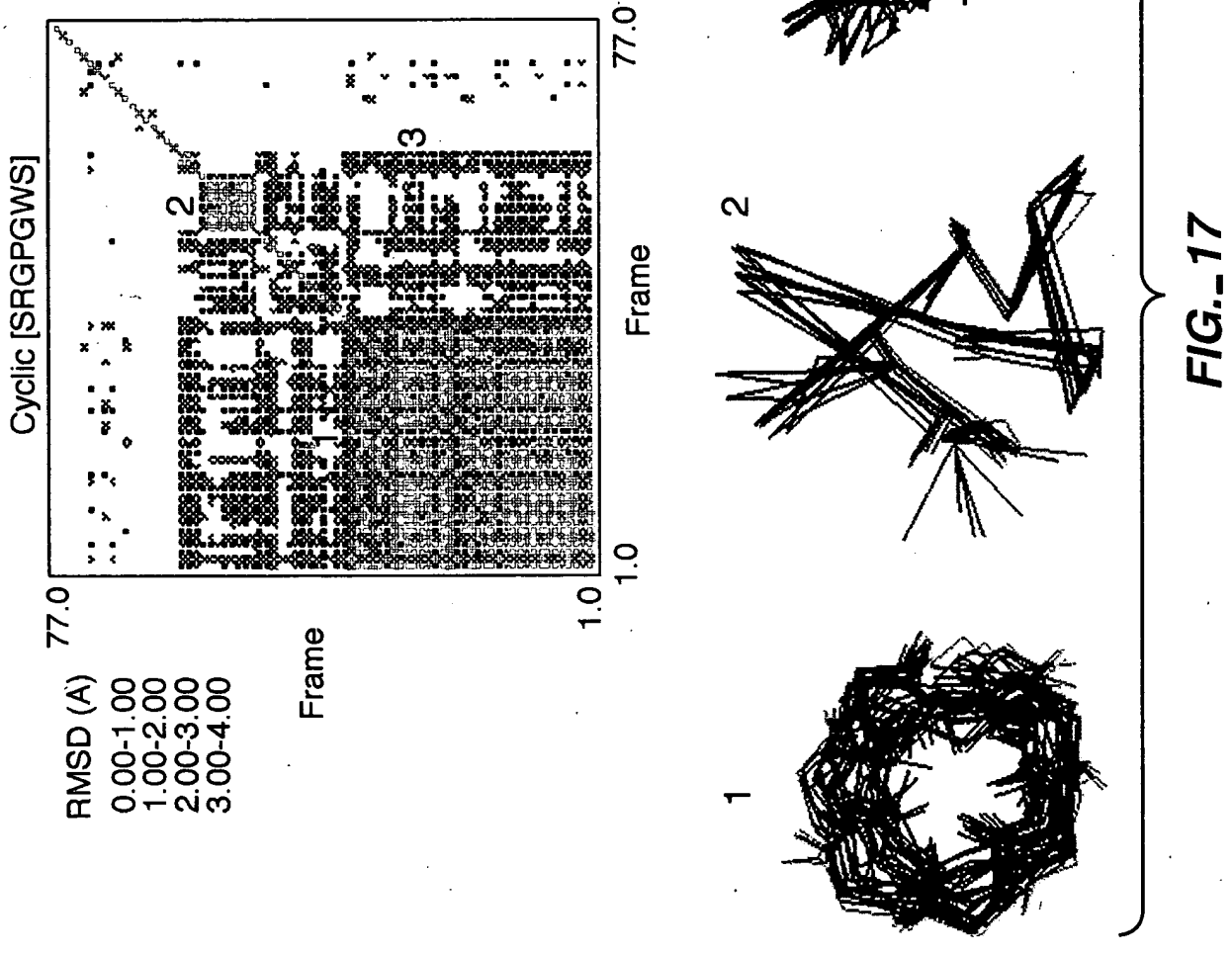

Cyclic [SRGPGWS]
RMSD (A)
0.00-1.00
1.00-2.00
2.00-3.00
3.00-4.00
77.0
1.0
Frame
1.0
77.0
Frame
1
2
3
1
2
3

FIG._17